수의학석사학위논문

영상 기반 기계학습을 이용한 돼지 폐병변의 진단

Diagnosis of swine lung lesion through image-based machine learning

2018 년 2 월

서울대학교 대학원

수의학과 임상수의학 전공

이 홍 석

영상 기반 기계학습을 이용한

돼지 폐병변의 진단

Diagnosis of swine lung lesion through

image-based machine learning

지도교수: 김 용 백 (D.V.M., Ph.D.)

이 논문을 수의학 석사학위논문으로 제출함

2017 년 11 월

서울대학교 대학원

수의학과 임상수의학 전공

이 홍 석

이홍석의 석사학위논문을 인준함

2017 년 12 월

위 원 장 김 대 용 (인)

부 위 원 장 김 용 백 (인)

위 원 류 덕 영 (인)

ABSTRACT

Diagnosis of swine lung lesion through image-based machine learning

(Supervisor: Yongbaek Kim, D.V.M., Ph.D.)

Hong-Seok Lee

Major in Veterinary Clinical Sciences (Veterinary Clinical Pathology)

Graduate School of Veterinary Medicine

Seoul National University

Slaughter check system has been applied to improve food hygiene and swine health schemes in the countries with advanced swine industry. Lung inspection is the most critical part of the slaughter check system. Recent advance in computer vision technology has led to the development of computer-aided diagnosis (CAD). As a pilot study prior to organ inspection using CAD, the correlation between gross lung scoring and pathologic diagnosis was investigated. Lung tissues and their gross images were collected from slaughterhouses. The images were subjected to gross lung lesion scoring. Histopathologic examination was conducted to classify

the lung lesions into bronchopneumonia and interstitial pneumonia. The gross lung lesion scoring revealed 100% of sensitivity and 77.3% of specificity for bronchopneumonia based on the 90% confidence interval. However, the specificity was low for the diagnosis of interstitial pneumonia. The area under receiver operation characteristic curve was 0.896, indicating a good discriminative performance of gross lung scoring for bronchopneumonia. Taken together, the data indicated that visual information of the photograph was useful to screen lung lesions. Further study was performed to establish a CAD model for swine pulmonary diseases. Lung tissues and the gross images were collected from the slaughterhouses and the lung lesions were histopathologically classified as bronchopneumonia, interstitial pneumonia, and pleural diseases. The *scale-invariant feature transform* algorithm was adopted to extract significant feature from the images. As a machine learning classification, *k-nearest neighbor* algorithm was applied to classify the extracted feature. For bronchopneumonia, the CAD model demonstrated the sensitivity of 96.7%, the specificity of 72.3 %, and accuracy of 82.0%. For interstitial pneumonia, the sensitivity was 75.8%, but the specificity and accuracy were high as 94.4% and 87.4%, respectively. However, it showed low performance for the diagnostic classification of pleural diseases. The present study provided a new approach of organ inspection through image-based machine learning, giving insight into application of CAD to slaughter check system. The data presented in this study could be a cornerstone for the development of computational image-based organ inspection system in veterinary diagnostics.

Keywords: Computer-aided diagnosis, Machine learning, Histopathology, Pneumonia, Slaughter check, Lung lesion scoring

Student number: 2017-21782

CONTENTS

ABBREVIATIONS

ANOVA: analysis of variance

AUC: area under curve

CAD: computer-aided diagnosis

CT: computer tomography

DOG: difference of Gaussian

FLANN: fast library for approximate nearest neighbors

FN: false negative

FP: false positive

ISO: international organization for standardization

KNN: k-nearest neighbor

PCR: polymerase chain reaction

PCV2: porcine circovirus type 2

PRRSV: porcine reproductive respiratory syndrome virus

RAM: random access memory

ROC: receiver operating characteristic curve

SD: standard deviation

SIFT: scale invariant feature transform

SRCD: swine respiratory complex disease

TN: true negative

TP: true positive

LITERATURE REVIEW

Introduction

Along with the innovation of computer machine with Moore's law, the computation speed has been accelerated (Strawn and Strawn 2015). As the speed of computation increases, attempts have been made to solve complex problem in the real world, and machine learning is a part of them (Fernández-Delgado, Cernadas et al. 2014, Lison 2015). In diagnostic field, the machine learning research is actively performed with imaging diagnostic technology such as x-ray, computer tomography, and magnetic resonance imaging (Sun, Wang et al. 2013, El-Dahshan, Mohsen et al. 2014, Lee and Chen 2015). Even within the massive CT image, it is possible to detect pulmonary nodules in a short time through deep learning which is a methodology of machine learning (Setio, Ciompi et al. 2016).

In the human medicine, not only imaging medicine, but also histopathology, cytopathology, and even internal medicine are researched with machine learning, referred to computer-aid diagnosis (CAD) (Kelly, Angelov et al. 2010, Fernández-Delgado, Cernadas et al. 2014). The CAD is an expert system that assists the diagnosis of medical personnel. There are four requirements as follows; CAD should enhance expert's performance; CAD should save time; CAD must be flexibly integrated into the workflow; CAD should not be held liability (van Ginneken, Schaefer-Prokop et al. 2011). Most CAD model in development do not

always meet above requirements. In principle, however, the above requirements must be met in order for the CAD model to be applied to clinical practice.

SRDCs in swine industry

Swine respiratory disease complex (SRDCs) is major cause of economic loss in world-wide swine industry (Chae 2005, Neumann, Kliebenstein et al. 2005). Despite SRDCs in growing-finishing pigs are usually asymptomatic, it decreases growth rate and industrial productivity (Chae 2005). It is challenging to distinguish between healthy pigs and SRDC pigs on premortem inspection. In countries with advanced swine industry, slaughter check system has been adopted to monitor swine diseases and establish swine health scheme for the industrial productivity (Baptista, Dahl et al. 2010, Jäger, McKinley et al. 2012). If total inspection is applied to slaughter check system, it could not only enhance the monitoring system and pig farm management, but also serve as a domestic surveillance system.

As a classical method of assessing the degree of swine pulmonary lesion, several lung score methods have been used (Madec and Kobisch 1982, Morrison, Hilley et al. 1985). In the previous study that assess pneumonia using lung scoring method, it is found that farm risk factor, carcass quality, and specific infectious agent is correlated with lung score (Ostanello, Dottori et al. 2007). In the present study, to identify the correlation between pulmonary diseases and the lung score obtained from visual information of digital image, hundreds of swine lung images were collected. As a follow-up study of a digital image with sufficient visual information

for pneumonia classification, numerical data were extracted from digital images and the CAD model based on machine learning that consisted of methodology in similar research (Burroni, Corona et al. 2004, André, Vercauteren et al. 2009).

CAD application into veterinary field

To the best knowledge of the author, unlike the CAD level already used in clinical practice through machine learning in the medical field, the study of applying machine learning is lacking in the field of veterinary medicine. To extrapolate the already developed CAD system of medical field to the veterinary field, it is required to consider various species characteristics and differences. Therefore, it is necessary not only to introduce the CAD system of the medical field but also to construct a veterinary CAD system that is based on the data collected in the veterinary field. Especially in radiology, because digital radiography is common in veterinary (Poteet 2008), and it is considered to be convenient for data collection that is labeled as bone fraction or pulmonary nodule, and the other radiological features. In the present study, gross image of swine lung was collected from slaughterhouses as background data to classify the detailed diagnosis of pulmonary lesion in swine.

In recent years, there has been numerous attempt to use computers on diagnostic field. The CAD models currently applied to the diagnostic field are based on raw data obtained through a professional diagnostic apparatus. Although numerous studies of CAD were progressively performed in diagnostic imaging field, the approach using gross images taken by non-diagnostic modalities is relatively scarce.

Because the purpose of this study is to develop an image-based CAD model for organ inspection, the gross images obtained through non-diagnostic modalities were subjected to machine learning techniques which have been applied in the previous studies. An CAD system was reported to assist in distinguishing between early melanoma and benign skin lesion through dermoscopic image which is similar to the gross image used in this study (Burroni, Corona et al. 2004). Other CAD studies which related to the gross image were endoscopy studies (André, Vercauteren et al. 2009, Liedlgruber and Uhl 2011).

Computer vision and machine learning technique applied to CAD

Because all pixel regions of image do not always have valuable information for detecting pulmonary lesions, it is a prerequisite to extract specific pixels which have significant difference compared to the peripheral pixels, termed as *feature* (Lowe 2004). In this study, the scale invariant feature transform (SIFT) was adopted for methodology of feature extraction, which has been widely used to recognize feature in various computer vision studies about lesion detection such as pulmonary nodule, bone fractions, and other mass on radiographic digital images (Awai, Murao et al. 2004, Kakeda, Moriya et al. 2004, Doi 2007). There are five values, which constitute one pixel. Two values indicate location of the pixel on the image, horizontal coordinate and vertical coordinate. Other three values have information of intensity for three colors which are red, blue, and green, however, there is just one intensity value in one pixel for grayscale image, so grayscale pixel has three

values. Each the intensity value ranges from 0 to 255. In SIFT method, entire pixel converted to grayscale. Also it is required to analyze differences of all pixel intensities among images which were several gradually blurred three times from the same image, using Gaussian blur methods, referred as to *Difference of Gaussian (DOG)* in research area of computer vision (Tsomko, Kim et al. 2010).

A classification model from numerical vectors, which are the extracted features of an image, is established through machine learning. Machine learning is a field of computer science which confers computers the ability to learn through training data set without precise programming (Samuel 2000). In pattern recognition, *k-nearest neighbor* (KNN) algorithm can be applied as non-parametric method used for classification (Khanmohammadi, Nasiri et al. 2007, Joglekar, Gedam et al. 2014). Every SIFT feature descriptors which are obtained from training images were plotted on the *feature space*, which is a 128-dimensional mathematical space. To classify features extracted from a testing image, a SIFT feature descriptor from a testing image is mapped into the *feature space,* and then the distances between the testing feature and trained features which are plotted at adjacent distances are calculated. To measure distance in 128-dimensional space, *Euclidean distance* can be adopted (Danielsson 1980). The basic principle of *k-nearest neighbor* is to predict the label of the testing feature through the most frequent label of the trained features (Altman 1992).

Purpose of the present study

With advances in computation processing power and the achievement of the datamining field, CAD has been widely studied and some of CAD model is already used in clinical practice. SRDC is one of the major causes inducing economic loss to the swine industry. Visual information obtained from swine pneumonia has been reported to be useful in classifying pulmonary diseases. As part of the swine health monitoring scheme, it is expected that CAD could be reinforce slaughter check system, however, CAD study applying machine learning is lacking in veterinary field.

In the present study, as a pilot study for establishment of image-based diagnostic system, the correlation between visual lung score and histopathologic diagnosis was assessed by lung scoring method, which has been widely used to record lung lesions (Madec and Kobisch 1982). Then the lung images were subjected to establishment of computer-aided diagnostic model, and diagnostic performance of the model was evaluated. Further, in this study, by analyzing the performance of the models for various training datasets, more suitable type of raw data and the number of raw data that is expected to be required for the target performance were suggested.

The aim of this study is to predict the histopathological diagnosis through only gross image without tissue sampling. The CAD model is still in prototype, but it is expected to improve the slaughter check system as an automated lesion monitoring system, and also a screening test that precedes histopathology.

CHAPTER 1.

EVALUATION OF CORRELATION BETWEEN GROSS LUNG SCORE AND MICROSCOPIC DIAGNOSIS FOR SWINE PNEUMONIA IN SLAUTHER HOUSES

Abstract

To reduce cost of swine production, a slaughter check system has been developed in countries with advanced swine industry. An evaluation of lung lesions in carcasses is a critical part of the slaughter check system. This study was performed as a pilot study for establishing image-based diagnostic system of swine lung. Lung tissues and their gross images were collected from slaughterhouses in Gyeonggi-do, South Korea. Scoring of the gross lung lesions was performed on the images of the lung. Histopathologic examination was conducted to classify the pulmonary lesions as bronchopneumonia and interstitial pneumonia. The lung lesion score was significantly higher in the bronchopneumonia group then in the interstitial pneumonia group ($p < 0.001$). A 90% confidence interval of the gloss lung lesion score was set for bronchopneumonia group, and it showed a sensitivity of 100% and specificity of 77.3%**.** The gloss lung lesion scoring test was subjected to the evaluation of diagnostic distinction by receiver operating characteristic curve, and appraised to have good discrimination for bronchopneumonia. An establishment of the gross lung lesion scoring test for the diagnosis of bronchopneumonia could be valuable for a screening test of macroscopic bronchopneumonia in the swine slaughter check system.

Introduction

Swine respiratory disease complex(SRDC) are frequently occurred in growing-finishing pigs, and cause economic loss through inducing death or inhibiting growth (Chae 2005, Neumann, Kliebenstein et al. 2005). Cause of SRDC is not only infectious agent such as bacteria and virus (Choi, Goyal et al. 2003), but also include environmental factor like as water supply and feeding system, hygiene status of farm, ventilation system, density of herd, temperature, and humidity (Hurnik, Dohoo et al. 1994, Stärk, Pfeiffer et al. 1998, Stärk 2000, Gardner, Willeberg et al. 2002). SRDC is classified as a upper respiratory disease and a lower respiratory disease (Robertson, Wilson et al. 1990, Chae 2005).

Pneumonia, the most common upper respiratory disease in slaughter, is classified into bronchopneumonia and interstitial pneumonia according to the pathogenesis (Hansen, Pors et al. 2010). Bronchopneumonia is commonly caused by bacterial infections such as *Mycoplasma hyopneumoniae* and *Actinobacillus pleuropneumoniae* (Redondo, Masot et al. 2009, Meyns, Van Steelant et al. 2011). Interstitial pneumonia is mainly caused by viral infection. Porcine circovirus type 2 (PCV2) and porcine reproductive respiratory syndrome virus (PRRSV) and the most common causative agents (Thacker, Halbur et al. 1999, Harms, Halbur et al. 2002, Hansen, Pors et al. 2010). If pneumonia type and causative agent could be estimated through visual inspection of slaughter check, it can be widely applied to improve management for individual farms. In countries with advanced swine industry, such as Sweden, Australia, and Denmark, result of the slaughter check is

used to manage the SRDC of the individual farms (Baptista, Dahl et al. 2010). However, in South Korea, there is relatively scarce effort to apply the results of slaughter check to the swine health scheme (Hwang and Han 2006). In domestic swine industry, scales of individual pig farm have become larger and the density of herds has been increased so that possibility of SRDC has risen (Kim, Hwang et al. 2011). Therefore, it is urgent to establish diagnostic criteria for respiratory diseases through slaughter check in the domestic swine industry and a farm managing system which utilizes the results of slaughter check to establish swine health scheme.

Madec and Kobisch(1982) proposed a method to assess the degree of gross lung lesions (Madec and Kobisch 1982). It is a universal method to grossly evaluate pulmonary lesions in most documents (Rueda, Bulnes et al. 2002, Vicca, Stakenborg et al. 2003, Sibila, Pieters et al. 2009). The lung scoring method is easier to be applied to slaughter check during slaughter process, because the evaluation is simple and scoring time is shorter than other evaluation methods such as calculating the weight for each lobe. In this method, each lobe was scored from 0 to 4 points based on the area ratio of consolidation. Visual lung score is the sum of points of seven lobes. Visual lung score ranges from 0 to 28 points (Madec and Kobisch 1982). Previous swine studies with visual lung lesion score were mainly concerned with clinical status of pigs, and environmental factor of farms (Wilson, Takov et al. 1986, Rueda, Bulnes et al. 2002, Vicca, Stakenborg et al. 2003, Feng, Shao et al. 2010). However, correlation between the visual lung lesion score and

histopathologic diagnosis is scarcely investigated. In the present study, swine lungs obtained from domestic slaughterhouses were investigated the visual lung lesion score. Analysis of the correlation between visual lung lesion score and histopathological diagnosis was performed and diagnostic criteria which is based on visual lung score is suggested. It is expected that visual lung scoring method can be applied to current slaughter check system to improve management of swine health scheme.

Materials and methods

Collection of tissue samples and gross images from swine lung

From November 2015 to April 2016, 66 disease lungs were subjected to visual inspection at a slaughterhouse in northern Gyeonggi-do, Republic of Korea. Lung lesions were photographed and surgically excised for histopathology. The gross images (2592 x 1944 pixels, 350 dots per inch) were photographed by digital camera (Sony, DSC-WX100). Each lung was photographed on front and back side. The surgically collected lung tissues were immediately fixed with 10% formalin solution.

Histopathological classification

A total of 66 swine lung lesions were classified into interstitial pneumonia and bronchopneumonia according to gross pathological classification criteria (Jeffrey, Locke et al. 2012). Tissues were collected from the gross lesion site for histopathology as confirmative diagnosis. When the different types of lesions were mixed, the tissues were taken for each lesion. In addition, the interstitial pneumonia is subdivided as mild, moderate, and severe proliferation of the alveolar epithelium.

Evaluation of visual lung score

The pulmonary lesions were identified by visual inspection of the lung images collected at the slaughterhouse. A simple lung scoring method was applied to the evaluation of individual lobes from each lung image (Madec and Kobisch 1982): 0 point if there is no lesion, 1 point if the lesion is found in less than 25% of the lobe, 2 point if the lesion is found in the region of 25-49% of the lobe, 3 points if lesions are detected in 50-74% area of the lobe, and when the lesion was found in more than 75% of the lobe area, it was evaluated as 4 points. According to the visual lung evaluation method, one lung composed of 7 lobes was evaluated from 0 to 28 points.

Statistical analysis

Statistical analysis was performed using Student's *t* test of SPSS Statistics (version 24; IBM Co., USA), for analysis of visual lung scores between interstitial pneumonia group and bronchopneumonia group classified by histopathological examination. For the comparative analysis between the subdivisions of mild, moderate and severe in interstitial pneumonia according to the degree of alveolar epithelium proliferation, one-way ANOVA of SPSS Statistics was used to analyze the significant differences among the subdivisions. The receiver operating characteristic curve (ROC curve) and the area under the curve (AUC) for the visual lung score between bronchopneumonia and interstitial pneumonia were obtained through python-based open source library for statistical analysis, scikit-learn (ver. 0.18.1).

Results

Visual inspection for primary classification of swine pneumonia

Visual inspection was conducted for primary classification of the 66 lung lesions as interstitial pneumonia and bronchopneumonia and the histopathological diagnosis was carried out for confirmative diagnosis. In case of interstitial pneumonia, lesions generally occurred over the entire area of the lung. Most of the lobes consistently were affected and showed a texture such as rubber, and it is noticeable that red color lesions were observed (Fig. 1-A). In the case of bronchopneumonia, most lesions showed hardened texture with dark reddish brown at the cranioventral side of lung, and the boundary between normal tissue and lesion site was clearly distinguished by interlobular septa. Exudate with white viscous fluid was observed during tissue excision (Fig. 1-B).

Histopathological examination for confirmative diagnosis

Histopathological examination of a total of 66 pig lung tissues revealed that interstitial pneumonia accounted for 69.70% of all pneumonia lesions (Table 1). In the case of interstitial pneumonia, the alveolar epithelium was proliferated more than the normal alveolar epithelium, not the monolayer, but was proliferated to several layers, or the alveolar wall thickened due to infiltration of neutrophils and alveolar macrophage. Bronchopneumonia accounted for 30.30% of all pneumonia

lesions, and neutrophil and alveolar macrophages were found in the alveoli and bronchioles (Fig. 2).

Interstitial pneumonia, the most diagnosed with 46 cases, was subdivided into mild, moderate and severe subgroups according to the hyperplasia of the alveolar wall. The mild interstitial pneumonia was identified in 14 cases, and accounted for 30.4% of the total interstitial pneumonia cases as the lowest rate. There were 15 cases of moderate interstitial pneumonia and 17 cases of severe interstitial pneumonia, accounting for 32.6% and 37.0% of the total interstitial pneumonia cases, respectively (Table2).

Analysis of visual lung score between histopathologically classified bronchopneumonia and interstitial pneumonia

For the 66 lungs collected, the visual lung score was evaluated according to the lung scoring method presented by Madec and Kobisch (1982) (Madec and Kobisch 1982). Student's *t*-test for visual lung score within the group according to histopathological examination revealed a significant difference in bronchopneumonia and interstitial pneumonia ($p < 0.001$). In order to reduce the outliers of visual lung score, samples corresponding to the maximum and minimum values of each class were excluded. For 18 samples of bronchial pneumonia and 44 samples of interstitial pneumonia, mean, median, and standard deviation were analyzed. In the case of bronchopneumonia, the mean value of the visual lung score was 21.15, which was 1.57 times higher than the mean value of visual lung score

in interstitial pneumonia, 13.43. Bronchopneumonia has a lower standard deviation than interstitial pneumonia. The distribution of visual lung scores in bronchopneumonia is denser than the distribution of visual lung scores interstitial pneumonia (Table 3) (Table 1 and 3).

Diagnostic assessment of visual lung score method for pneumonia classification

The true positive, true negative, false positive, false negative, sensitivity, and specificity were analyzed after setting 95% and 90% confidence intervals from the mean and standard deviation of the visual lung score in each group of bronchopneumonia and interstitial pneumonia. For bronchopneumonia group, the sensitivity was 100% and the specificity was 77.3% in the 90% confidence interval, which has higher specificity than the 95% confidence interval (Table 4). I n the case of the visual lung scores of the interstitial pneumonia group, sensitivity was high as 97.7% in 95% confidence interval, but the specificity was low as 27.8%. For 90% confidence interval of the visual lung score in the interstitial pneumonia group, the sensitivity decreased to 93.2% and the specificity increased to 38.9% compared to the diagnostic indicators of 95% confidence interval. In 90% confidence interval of visual lung score in interstitial pneumonia, the sensitivity slightly decreased and the specificity increased more than the decrease in sensitivity compared in 95% confidence interval of the visual lung score, but still the specificity was low.

Histopathology was conducted as the gold standard for classification of

bronchopneumonia and interstitial pneumonia. For bronchopneumonia, performance of discrimination was assessed by plotting the receiver operating characteristic curve (ROC curve) of the visual lung score. Performance evaluation was carried out by calculating the area under the curve (AUC) of the ROC curve. The AUC of the ROC curve for visual lung score in bronchopneumonia was calculated as 0.8958 which indicates the visual lung scoring is useful method for bronchopneumonia diagnosis (Fig. 4).

Interstitial pneumonia was divided into subgroups of mild, moderate, and severe according to the degree of thickening of the alveolar wall by an additional microscopic examination. After excluding the samples with the maximum and minimum values in each subgroup to reduce the outliers, the mean, median, and standard deviation were examined and the visual lung score in each subgroup were analyzed. The same median value was found in the visual lung scores of subgroups of interstitial pneumonia classified as mild and severe. The highest difference among the three groups was 2.5 in median, and 0.33 in standard deviation. One-way ANOVA for the visual lung score of the three subgroups of interstitial pneumonia did not reveal a statistically significant difference (p = 0.407). The sensitivity and specificity were found to be low regardless of the three subgroups of the interstitial pneumonia or the confidence rates of 95% and 90% in interstitial pneumonia (data not shown).

Table 1. Incidence and gross lung lesion score of pneumonia diagnosed by histopathology in slaughtered pigs

	Incidence	Mean	Median	SD
Interstitial pneumonia	46 (69.70%)	13.43	13	5.25
Bronchopneumonia	20 (30.30%)	21.15	21	3.31

Table 2. Incidence and gross lung lesion score of interstitial pneumonia with different severity

	Incidence (proportion)	Mean	Median	SD
Mild interstitial pneumonia	14 (30.4%)	11.9	13	4.41
Moderate interstitial pneumonia	15 (32.6%)	13.5	16	4.52
Marked interstitial pneumonia	17 (37.0)	14.4	13	4.74

Table 3. Interval frequency distribution of gross lung lesion scores in bronchopneumonia and interstitial pneumonia groups

	Interstitial pneumonia	Bronchopneumonia	Total
0 - 7	8	0	8
8 - 14	21	1	22
15 - 21	12	10	22
22 - 28	5	9	14

Table 4. Ranges of gross lung lesion scores with 95% or 90% of confidence rate, and their sensitivities and specificities

	Confidence rate	Lower limit	Upper limit	TP	TN	FP	FN	Sensitivity	Specificity
Broncho-	95%	16.54	26.24	18	31	13	0	1.000	0.705
pneumonia	90%	17.32	25.46	18	34	10	0	1.000	0.773
Interstitial	95%	3.15	23.71	43	5	13	1	0.977	0.278
pneumonia	90%	4.80	22.06	41	7	11	3	0.932	0.389

TP: true positive; TN: true negative; FP: false positive; FN: false negative

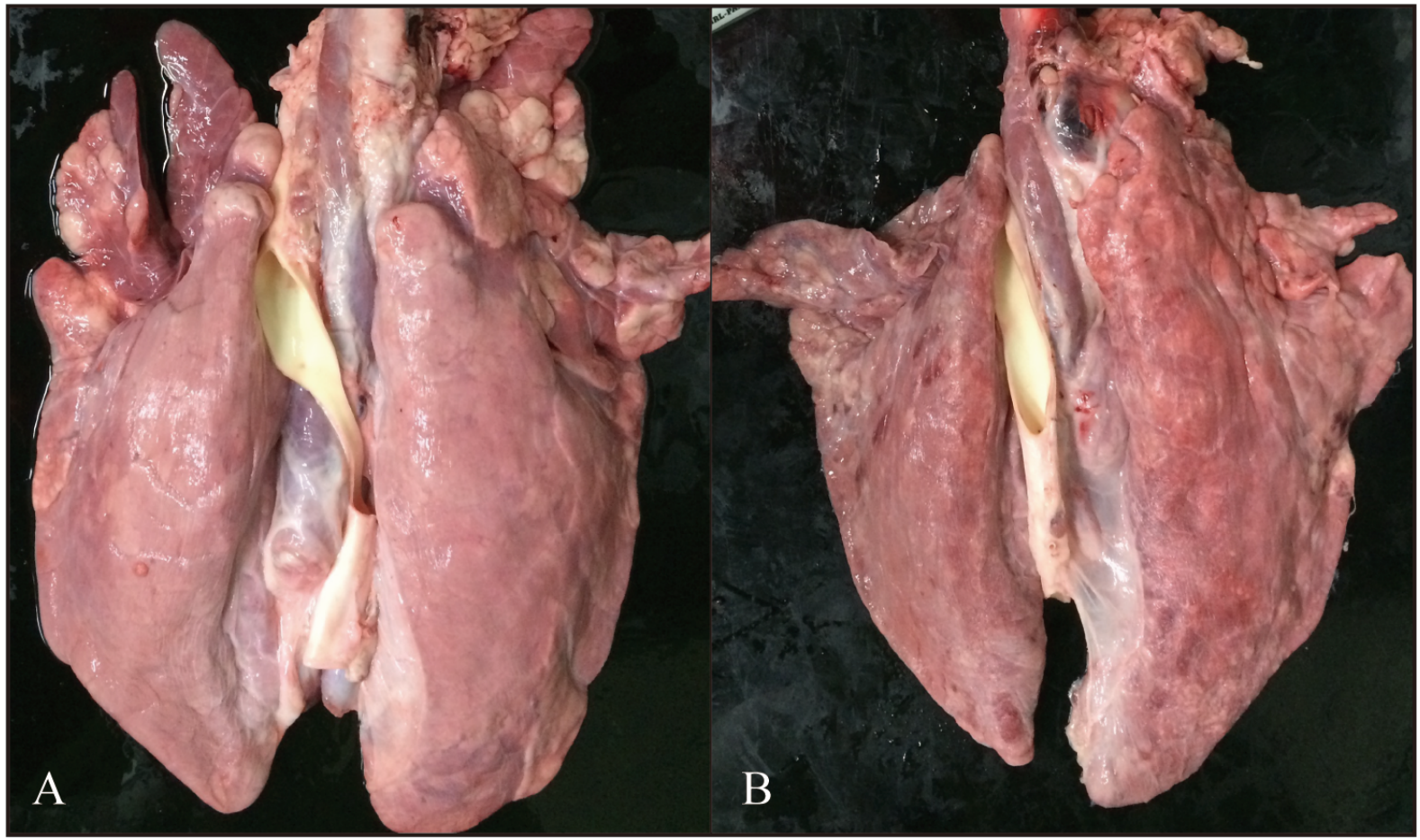

Fig. 1. Representative gross image of swine pneumonia; (A) Bronchopneumonia, (B) interstitial pneumonia.

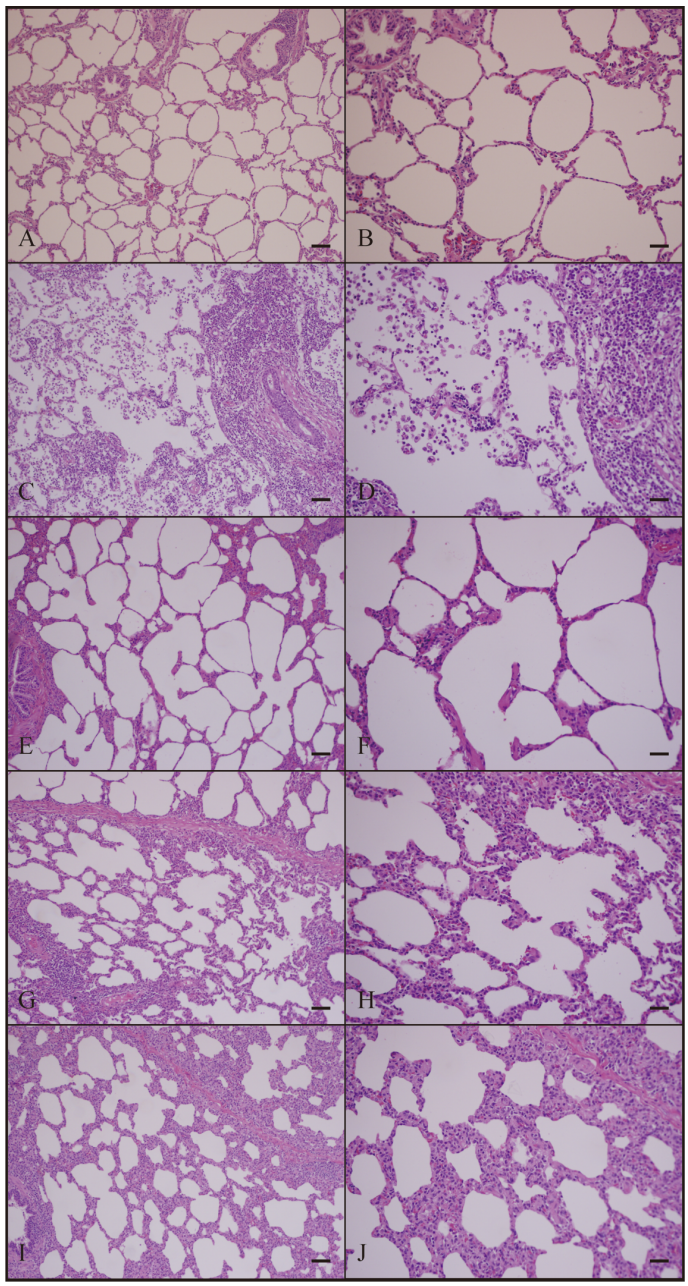

Fig. 2. Microscopic images of lung sections stained with hematoxylin and eosin; (A and B) normal lung, (C and D) bronchopneumonia, (E and F) mild interstitial pneumonia, (G and H) moderate interstitial pneumonia, (I and J) severe interstitial pneumonia. Scale bars = 50 um (B, D, F, H, and J) or 100um (A, C, E, G, and I).

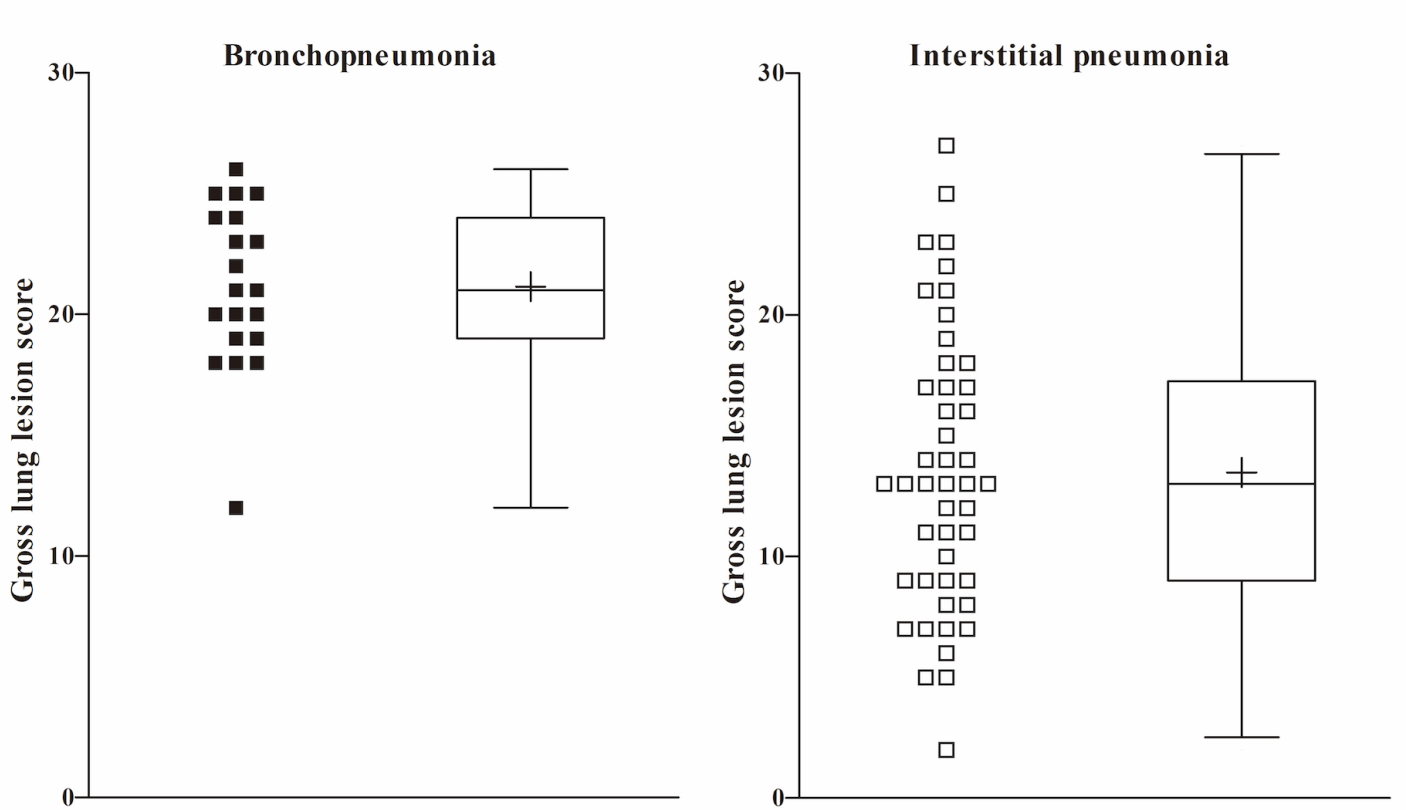

Fig. 3. Distribution of gross lung lesion score and 90% confidence intervals of bronchopneumonia and interstitial pneumonia groups.

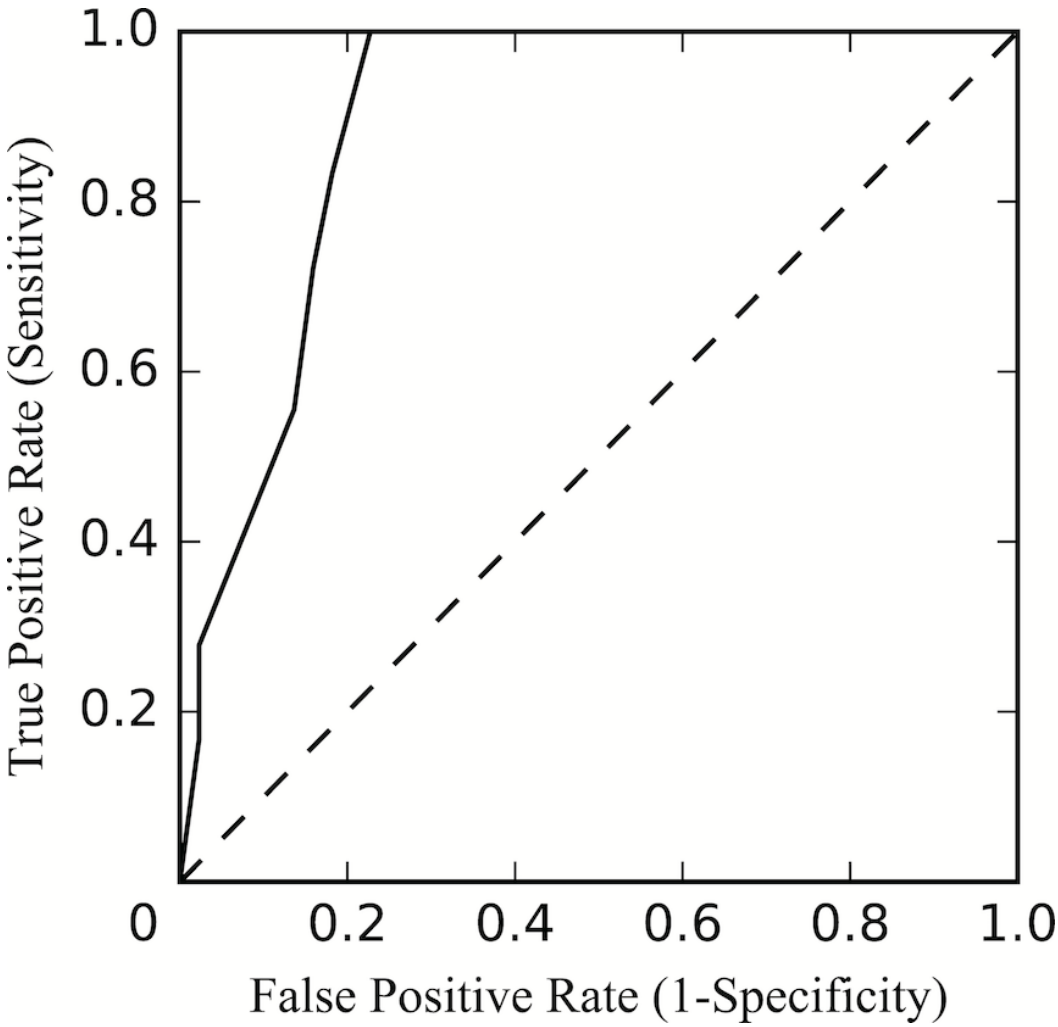

Fig. 4. Receiver operating characteristic curve of the gross lung lesion score in bronchopneumonia group. An area of under the receiver operating characteristic curve is 0.8958.

Discussion

Swine disease monitoring through slaughter check, which is requisite for establishment of swine health scheme and food hygiene has been widely carried out in countries with advanced swine industry (Mousing, Jensen et al. 1997). Swine disease monitoring through slaughter check has been applied to investigate prevalence of specific disease, to establish policy, and to manage swine farms with pig tracking system (Christensen, Ellegaard et al. 1994). Especially in Sweden, Denmark, and Australia, the slaughter check systems are well established as a swine disease monitoring system for finishing pigs from each farm (Hwang and Han 2006). Generally, serologic test and polymerase chain reaction (PCR) test are required to check the infectious disease (Thacker 2004). In domestic swine industry, however, serologic tests and genetic tests are difficult to apply widely slaughter check because of problems such as sample collection, sample transportation, and cost of these tests. In contrast, evaluation of the visual lung score during slaughter check process is applicatory to domestic slaughterhouses because the scoring method is simple and the cost is much less than serological and genetic tests.

In the present study, the evaluation of visual lung score was carried out through visual inspection on the digital images of swine lung. In previous study, Pearson correlation analysis showed that there is an appropriate correlation ($r = 0.94$) between lung scoring accompanied by palpation test and gross lung scoring excluding palpation test (Hurnik, Hanna et al. 1993). It is considered that the visual lung scoring is applicatory to slaughter check because this method can be conducted

simply through the images which were photographed during slaughter process. In addition, it has been shown in previous studies that the agreement between the two-dimensional evaluation and the three-dimensional evaluation of lung lesions was high (Hurnik, Hanna et al. 1993, Davies, Bahnson et al. 1995). Therefore, visual lung score method based on two-dimensional images is considered to be valid for evaluation of swine lung lesions. Furthermore, the evaluation of visual lung score from lung images was considered to be applicable to assess the correlation with histopathological classification.

The lung scoring method proposed by Madec and Kobisch (1982) has been a widely used to evaluate lung score for recording swine pneumonia lesions (Rueda, Bulnes et al. 2002, Vicca, Stakenborg et al. 2003, Sibila, Pieters et al. 2009). Since the lung scoring method is simpler than other scoring methods which calculate by different weighting of each lobe, it is more suitable for application to the current slaughter process. Previous studies on the gross lung score and swine pneumonia revealed a correlation between the score and enzootic pneumonia (Huhn 1970), and the sensitivity and specificity of the scoring classification were good for enzootic pneumonia (Hurnik, Hanna et al. 1993).

In the present study, lung tissues of slaughtered pigs were histopathologically classified into bronchopneumonia and interstitial pneumonia, and the visual lung scores between the two classes were statistically different ($p < 0.001$). The diagnostic value of the visual lung scoring was low in the diagnosis of the interstitial pneumonia. On the other hand, the 90% confidence interval of the visual lung score

in the group of bronchopneumonia showed good sensitivity and moderate specificity, so that the visual lung scoring is expected as a screening test for bronchopneumonia. The visual lung scoring can be a good screening test because the mean visual lung score in the group of bronchopneumonia is significantly higher than that in the interstitial pneumonia group, and the standard deviation in bronchopneumonia group is low so that the values are densely distributed near the mean value statistically.

In this study, the receiver operation characteristic curve (ROC curve) was analyzed to evaluate the diagnostic discrimination of visual lung score for bronchopneumonia. The ROC curve shows the change in sensitivity and specificity according to the cut-off values. The area under curve (AUC) value of ROC curve ranges from 0.5 to 1. The higher the AUC value, the better the diagnostic discrimination is considered. In general, an AUC value of 0.5 is considered to have no diagnostic value, an AUC value between 0.5 and 0.7 was not significantly better than the random discrimination. An acceptable discrimination is indicated when the AUC value is 0.7 to 0.8, an AUC value between 0.8 and 0.9 suggests an excellent discrimination, and a diagnostic model with outstanding discrimination is suggested when the AUC value is greater than 0.9 (Hosmer Jr, Lemeshow et al. 2013). Histopathology was applied as the gold standard, then the discrimination of visual lung scoring was investigated for bronchopneumonia. The AUC value of the ROC curve was 0.8958, which indicates that the visual lung scoring has excellent discrimination in distinguishing between bronchopneumonia and non-

bronchogenic pneumonia.

There was significant difference in the visual lung score between the groups of interstitial pneumonia and bronchopneumonia. The visual lung scoring was evaluated as an excellent method for distinguishing bronchopneumonia. Bacterial pathogens have been considered as main causative agents of bronchopneumonia, and the higher the visual lung score, the higher the likelihood of detection of bacterial pathogens (Lee, Hwang et al. 2011). It is considered that bacterial pneumonia can be estimated by screening test for bronchopneumonia through evaluation of visual lung score. In contrast, the confidence interval of the visual lung score in the interstitial pneumonia group has low diagnostic competence because of the low specificity. In addition, one-way ANOVA indicated no significant difference in visual lung scores among the three subdivisions of interstitial pneumonia (p = 0.4069). Therefore, a diagnostic approach based on visual lung score is not useful for distinguishing the subdivisions of interstitial pneumonia.

In the present study, the swine lungs were diagnosed by histopathology and the correlation between visual lung score and histopathological diagnosis was revealed. Then the cut-off value of the visual lung score for screening bronchopneumonia was secured through diagnostic evaluation. As a screening test for bronchopneumonia, the 90% confidence interval, which has 100% of sensitivity and the highest specificity among the confidence intervals ranges 18 to 25 points. Although the differential diagnosis of pathogenic agent is limited only with

diagnosis of bronchopneumonia, the visual lung score was found to be a diagnostic index of excellent discrimination based on the ROC curve for bronchopneumonia, and previous studies have shown the correlation between bacterial pathogenicity and the lung score (Lee, Hwang et al. 2011, Hosmer Jr, Lemeshow et al. 2013). However, the diagnostic approach of the lung score was limited to confirm the correlation. Further, in this study, the cut-off interval of the visual lung score for bronchogenic pneumonia was elucidated as a screening test with excellent sensitivity and moderate specificity, suggesting pathogen to examine first. This is a new approach to quantitative prediction of histopathological classification through visual lung scoring, and the scoring method can be applied as a simple and useful screening test for bronchopneumonia in the diagnostic field.

CHAPTER 2.

CLASSIFICATION OF SWINE LUNG LESIONS BY A COMPUTER-AIDED DIAGNOSTIC MODEL CREATED USING IMAGE-BASED MACHINE LEARNING ALGORITHM

Abstract

Advancement in computer vision technology has facilitated the progress of the computer-aided diagnosis (CAD). The present study was performed to develop a diagnostic system based on an image analysis of swine lung lesions and to assess its diagnostic performance. Using gross and histopathologic examinations, the lung lesions were classified into bronchopneumonia, interstitial pneumonia, pleuropneumonia and pleuritis. From the digital images of the gross lung lesions, the computerized features were extracted using the *Scale-invariant feature transforms,* a methodology of image analysis. The *k-nearest neighbor* was adopted as a classification algorithm of machine learning. Multiple CAD models were generated using various training sets that were organized based on the characteristics of images and group size. From each training set, a randomly selected 10% testing subsets was used to verify the reliability of the models. A training set that consists of both complete and close-up lung images showed the best performance among the other training sets. Classification performance of the model for bronchopneumonia was sensitivity of 96.7%, specificity of 72.3% and accuracy of 82.0%. For classification of interstitial pneumonia, the model had specificity of 94.4%, sensitivity of 75.8% and accuracy of 87.4%. This is the first study adopting image-based machine learning algorithm for inspecting animal organs. Although still a prototype, the data presented in this study provides insight into the applicability of CAD to slaughter check system that predicts histopathological diagnosis, and gives a direction for further study to enhance the performance of image-based organ inspection.

Introduction

Swine respiratory disease complex (SRDC) has been a leading cause of economic loss in world-wide swine industry (Chae, 2005; Neumann et al., 2005). Despite SRDC of growing-finishing pigs is usually asymptomatic, it impinges growth rate and productivity in swine industry (Chae, 2005). Except severe cases of SRDC, it is challenging to distinguish between healthy pigs and pigs with SRDC by ante-mortem inspection. In countries with advanced swine industry, slaughter check system has been adopted to monitor health status of swine herd and to enhance the productivity. Furthermore, not only for just monitoring, the slaughter check system has been also used to establish swine health scheme (Baptista et al., 2010; Jäger et al., 2012). A regular slaughter check for finishing pigs could be used to assess the effectiveness of farm management and disease surveillance and control.

The slaughter check system consists of lesion monitoring and seromonitoring. The lesion monitoring targets endemic diseases incurring large economic loss such as enzootic pneumonia, pleuropneumonia, atrophic rhinitis, liver white spot, and other inflammatory or infectious lesions (Meyns et al., 2011; Sanchez-Vazquez et al., 2011). The seromonitoring is for monitoring infectious diseases such as Aujeszky's disease, porcine respiratory reproductive syndrome, hog cholera, and swine influenza (Elbers et al., 2000; Maes et al., 1999). One of requisition for slaughter check is applicability to field condition of slaughterhouse. The slaughter check should be carried out with minimal interruption to production process in slaughterhouse. Considering the applicability, lesion monitoring is more accessible

approach than seromonitoring that requires blood collection, shipping, and hemoserological test. For the lesion monitoring, only simple palpation and visual inspection are necessary during the slaughter process.

For the lesion monitoring, lung inspection is one of the most critical parts. Swine pulmonary lesions are generally classified as bronchopneumonia, interstitial pneumonia, pleuritis and pleuropneumonia. Because the inspection is carried out by human inspectors, the cost is enormous when all pigs are subjected to the slaughter check system. Currently, slaughter check system is performed on a subset of randomly selected swine herds, which has limitation of partial and inconsistent evaluation. With regard to health monitoring scheme, continuous slaughter check makes more sensitive surveillance of swine diseases, and it also could facilitate the swine disease control.

With the recent improvement in computer vision technology, computer-aided diagnosis (CAD) has been applied to the field of diagnostic imaging such as radiography, computed tomography, and magnetic resonance imaging (El-Dahshan et al., 2014; Gurcan et al., 2002; Sun et al., 2013). Although numerous studies about CAD were performed in diagnostic imaging area, the approach using gross images taken by non-diagnostic modalities is relatively scarce. CAD model based on gross lung images which are photographed using common digital camera has not been developed. If a CAD can be adopted in slaughter check system, cost of health monitoring scheme would be significantly reduced. Moreover, inspection of all pigs

could be performed at the slaughterhouse, enhancing food hygiene and disease surveillance.

To generate CAD model based on image-based machine learning, methodologies for *feature* extraction and classification model are prerequisites. All pixels of an image are not always useful. For instance, background region that has no information about the lung and even has bloodstained area prevents an effective analysis. Therefore, it is essential to extract pixels that have meaningful information, referred to *feature*. Scale invariant feature transform (SIFT) has been widely used to recognize the features in many computer vision studies that aimed to detect lesions on radiographic digital images, such as pulmonary nodules, bone fractions, and other masses (Awai et al., 2004; Doi, 2007; Kakeda et al., 2004).

A classification model is established using the extracted features and label that is made by histopathology. The *features* with the correspondent label constituted basic information working as a classification criterion that is referred to *training*. Using SIFT as *feature* extractor and k-nearest neighbor (KNN) classification (Jiang et al., 2015; Kashif et al., 2016; Shouman et al., 2012), the present study established a CAD model for classifying gross images of swine lung lesions. Because the CAD model could be utilized for classifying swine lungs in the slaughterhouse with minimal equipment, it is anticipated that the adoption of the CAD model into the slaughter check system enables lung inspection of all pigs at low cost, improving food hygiene and pig health monitoring scheme.

Materials and Methods

Collection of gross images from swine lung

From slaughterhouses, 773 of gross lung images and lung tissues corresponding to each image were collected. The pulmonary lesions were photographed by digital camera (Sony, DSC-WX100, 2592 x 1944 pixels, 350 dots per inch). In order to maintain consistency, the subjects were photographed at constant distance with fixed exposure time (0.067s, 1/15), aperture (f/3.3), and international standards organization setting (800).

Four-category classification of swine lungs based on gross and histopathology examination

The swine lung lesions were classified into bronchopneumonia, interstitial pneumonia, pleuropneumonia and pleuritis, based on gross and histopathological findings (Jeffrey et al., 2012). For histopathology as a confirmative diagnosis, representative areas of each lung were excised and immediately fixed in 10% normal buffered formalin solution for at least 2 days. After fixation, the lung tissues were routinely processed for embedding in paraffin. The paraffin blocks were cut into sections with 3-4 μm thickness using a microtome and mounted on glass slides. After staining with hematoxylin and eosin, all slides were examined using a light microscope. The gross images of lung lesions were labeled using the histopathology data. When two or more diseases were diagnosed in one lung tissue, the image was labeled as one of more dominant diseases.

Feature extraction from the gross images using scale invariant feature transform (SIFT)

OpenCV, an open source library of computer vision with a good applicability, was adopted to analyze the digital images of swine lungs (Aly et al., 2011). Since all pixels of entire region did not always have meaningful information for detecting pulmonary lesions, it was required to extract significant pixels termed as *feature* by calculating differences from peripheral pixels (Lowe, 2004).

In this study, the SIFT was adopted as a methodology for feature extraction (Awai et al., 2004; Doi, 2007; Kakeda et al., 2004). Each pixel of gray-scaled image was expressed as one intensity value. Entire pixels of an image were converted to grayscale and intensity differences of each pixel were calculated from neighbor pixels and gradually blurred pixels using Gaussian blur methods (Tsomko et al., 2010). By analyzing the difference of the pixels, a specific *keypoint* was obtained based on the point where the intensity of the pixel had 10 or more difference from that of the surrounding and gradually blurred pixels (Fig. 1). The SIFT feature descriptor was extracted by calculating the differences of intensity between the *keypoint* and each of 128 pixels of square area centered on the *keypoint*. Each SIFT descriptor had 128 numeric values, referred as 128-dimentional vector.

Organizing gross images to evaluate the influence of image characteristics and quantity on the diagnostic performance

To evaluate the effect of image characteristics on classification performance, gross images of swine lung were organized into three groups; complete lung group, close-up image group, blended image group. The complete lung group consisted of photographs showing all areas of the lung on each image. The close-up image group had only a part of a lung with close-up of lung lesions. The blended image group was a union of groups of complete lung and close-up images that blended in equal proportion as same size of complete lung group (Table 2). In order to determine the influence of image quantity on classification performance, four groups with different number of images in the groups were randomly organized; groups with one-fourth, two-fourth, three-fourth and total number of images (Table 3).

To evaluate influence of image characteristics and quantity, each image group was further divided into training and testing subsets. Randomly selected 90% of images in each group were assigned as a training subset, and the remained 10% in each group were assigned as a testing subset. In order to avoid random deflection, testing subsets were made by organizing images of bronchopneumonia, interstitial pneumonia, pleuritis and pleuropneumonia into equal proportions. Five different testing subsets for each model were randomly selected and subjected to assessment of the diagnostic model, which is expressed as average value of sensitivity, specificity and accuracy.

Feature classification with k-nearest neighborhood (KNN)

By classifying the SIFT descriptors of swine lung image using KNN classification (Joglekar et al., 2014; Khanmohammadi et al., 2007), CAD models for swine pulmonary diseases were established. To create a classification model with KNN for 128-dimensional vectors, fast library for approximate nearest neighbors (FLANN) was applied (Muja and Lowe, 2009). Two groups of the SIFT feature descriptors were extracted from training subset and testing subset, respectively. All SIFT feature descriptors of the training subset were scattered in 128-dimensional *feature space.* Then, the SIFT feature descriptors from one image of the testing subset were mapped into the feature space. From existing features in the 128-dimensional space, distances to newly extracted feature of testing image were calculated. Among the existing features, two nearest neighbors were selected for newly extracted feature derived from the test image.

To measure distance between two features on the *feature space*, *Euclidean distance* was adopted, which is calculated as a square root of sum of the squares of the difference of each element between two different features (Weinberger et al., 2006). *Euclidean distances* between the newly mapped testing feature and the existing trained features from the training subset were calculated to select two trained features as *neighbor* which were nearest two to the testing feature (Fig 1). If the ratio of two *neighbors* was less than 0.8, the nearest one was adopted as a valid *neighbor* to avoid overfitting (Hawkins, 2004). After calculating the valid *neighbors* from all features of the test image, the diagnostic label of the training

image, where each *neighbor* was derived, was counted. The classification of each test image was predicted by the diagnostic label with the most *neighbor* count.

Data and statistical analysis

Descriptive statistics were applied (mean, standard deviation). Euclidean distance was applied to calculate the distance of the 128-dimensional vector (Weinberger et al., 2006). Sensitivity, specificity, and accuracy for bronchopneumonia, interstitial pneumonia, pleuropneumonia, pleuritis were determined as average of 5 trials with random sampling of test set for evaluating performance of multi-classification (Sokolova and Lapalme, 2009). Pearson correlation coefficient was used to assess the correlation between performance indicator and quantity of training data (Cohen, 1988). Measurement of image area was performed using ImageJ software (ImageJ, US National Institutes of Health, Bethesda, MD; imagej.nih.gov/ij/).

Results

Swine lungs were classified into four classes based on gross and histopathology.

Total number of the swine lung images was 773 which were composed of two groups, one with all area of entire lung and the other with partial close-up area of lung (Table 1). Based on the gross and microscopic findings, the lung lesions were labeled as bronchopneumonia, interstitial pneumonia, pleuritis, and pleuropneumonia. Representative gross images of bronchopneumonia, interstitial pneumonia, pleuritis and pleuropneumonia were shown in Figure 2. Background of the images was mostly white color with contaminated blood.

Lungs labeled as bronchopneumonia had dark, red-brown areas in cranioventral lobe, and the boundary between normal tissue and lesion site was clearly delineated by interlobular septa. Inflammatory cells including neutrophils and alveolar macrophages were present in the alveolar and bronchiolar spaces. Lesions of the interstitial pneumonia were generally present diffusely over the entire lung with rubbery texture. In histopathology, interstitial pneumonia showed thick alveolar septae by proliferation of alveolar epithelium and infiltration of inflammatory cells. In case of pleuritis, fibrous adhesions between the lung and pleural wall were identified with minimal lung lesions. In histopathology of pleuritis, serosa was thickened by fibrinous material with neutrophils and macrophages. Pleuropneumonia was diagnosed by fibrous adhesions between the thoracic wall and pleural surface of affected lung, and presence of the lung lesions at caudal lobes

with occasional hemorrhage. In histopathology, pleuropneumonia had multifocal areas of necrosis with infiltration of neutrophils and alveolar macrophages in bronchi and bronchioles.

SIFT extractor and KNN classification showed robustness to blood contaminated image.

To investigate the suitability of feature extraction to examine lung image, extracted features were visualized on gross lung image. In most gross images, blood contamination and bubbles were identified in background, but the most features were located in lung area. In representative of complete lung image (Fig. 3-A), 532 of features were identified. Among the extracted feature, features located outside of lung were 44, which is 8.27% of total extracted features. For close-up lung image (Fig. 3-B), 516 features were extracted in representative image, and the feature located outside of lung were 15, accounting for 2.90% of total extracted feature. In figure 3-A, lung area was measured as 64.06% of entire image area. Although 35.04% of image area was blood contaminated, 91.73% of features were extracted from lung area, which is region of interest. Blood contaminated area of figure 3-B was 29.97% of entire pixels. Although most background was contaminated by blood, 97.10% of features correctly extracted from close-up lung area.

From two images, the extracted SIFT features were visualized as colored circles on the gross lung images (Fig. 3). One image contained all area of lung and the other image had partial close-up of the same lung. Most features in the image were

located at the borderline and in lesions of the lungs, and those on the areas of backgrounds with blood contamination were minimal. According to KNN, couples of nearest *neighbor* features in feature space were visualized as colored lines. (Fig. 3). Corresponding lesions between the two images showed connectivity through a straight line connecting the two neighbor features on the visualized image (Fig. 3). Although several misconnections were shown as neighbor features of background to background and background to lung area, the large proportion of connected neighbor couple was lining between features on the lung region consisting evidence of decision making.

The blended image group showed better performance to detect pleuritis than the other groups.

In order to evaluate the influence of image characteristics on model performance, images were divided into three groups which have analogous size; complete lung group (n = 386), close-up lung group (n = 387), and equally blended group of formers (n = 386). Diagnostic performance of each CAD model was determined as an average of five trials with randomly selected testing subset (Table 2).

For bronchopneumonia, the sensitivity was highest in blended lung group as 93.3%. The specificities were 80.2% in complete lung group and 64.0% in blended group. In case of interstitial pneumonia, the sensitivities were 85.6% in complete lung group and 68.7% in blended group. The highest specificity was identified as 98.4% in close-up lung group. Except for close-up lung group, blended group

showed highest specificity of interstitial pneumonia, as 85.4%. In close-up lung group, however, diagnostic indicator showed bias to sensitivity or specificity. For every diagnosis, either sensitivity or specificity was lower than 50% in close-up lung group.

For pleuritis, the sensitivities were 0% in both groups of complete and close-up lungs. In contrast that both groups of complete and close-up lung could not detected pleuritis, the equally blended group showed 29.5% of sensitivity for pleuritis that was indicating an ability to detect pleuritis. Because testing subset of pleuropneumonia in 380s of lung image was just one or two, which was insufficient number to evaluate performance of classification, so the evaluation for classification of pleuropneumonia was not available (data not shown).

Classification performance showed rising by increasing number of training images.

For The model established using total image (n = 773), the classification of bronchopneumonia showed 96.7% of sensitivity, 82.0% of accuracy, and 72.3% of specificity. In classification of interstitial pneumonia, 94.4% of specificity and 87.4% of accuracy was identified, and the sensitivity was 75.8%. About pleuritis, sensitivity was low as 63.6%, but the specificity and accuracy were 98.9% and 92.8%, respectively.

In order to investigate classification performance by quantity of training data, four groups were randomly generated from total training images; groups with one-

fourth, two-fourth, three-fourth, total number of lung (Table 3). The investigation was performed five times to figure out the average values. In bronchopneumonia, the specificity and accuracy showed strong correlations with the quantity of training data in analysis of Pearson's correlation ($r = 0.806$, $p < 0.001$, and $r = 0.841$, $p < 0.001$, respectively). Although sensitivity of bronchopneumonia did not show significant correlation with data quantity, its sensitivity was high in any groups (Fig. 4). The correlation between the specificity of interstitial pneumonia and data quantity was not significant, but the specificity of interstitial pneumonia was high even in quarter-size group. (Fig. 4). The sensitivity and accuracy of interstitial pneumonia were showed high correlations with the number of training data ($r = 0.791$, $p < 0.001$ and $r = 0.788$, $p < 0.001$, respectively). About pleuritis, the sensitivity showed a highly strong positive correlation with the data quantity ($r = 0.892$, $p < 0.001$), and the accuracy had a positive correlation with quantity of data ($r = 0.620$, $p = 0.004$). Although significant correlation was not identified between its specificity and training data quantity, specificity was high even in the first quartile group. (Fig. 4). In pleuropneumonia, the sensitivity was very low, and the specificity was biased (data not shown). However, a positive correlation between its sensitivity and the number of training data was identified ($r = 0.484$, $p = 0.030$).

Individual operation time was increased according to enlarged group size of the training data.

Based on the KNN algorithm, the individual operation time which was required to perform classification was investigated. To investigate variation of operation time according to the number of images in the model, four groups of a quarter, a half, three-quarter, and entire number of images were examined. (3.3). The testing environment was Intel CPU G2020 2.90 GHz, 6.00 GB RAM, 256 GB hard disk drive, and 64-bit Windows operating system. As the number of image in the model increases, the averages of the individual operation time were gradually increased (Fig. 5). The mean values of individual runtime for a quarter, half, three quarter, and entire images were 1738, 1795, 1891, 1994 milliseconds, respectively.

Table 1. Number of labeled gross images of swine pulmonary lesion according to visual inspection and histopathological examination

Label	Images of close-up lung	Images of complete lung
Bronchopneumonia	188	188
Interstitial pneumonia	94	95
Pleuropneumonia	22	22
Pleuritis	82	82
Total	386	387

Table 2. Evaluation of swine lung classification for three groups based on image characteristics

	Type of training data	Sensitivity	Specificity	Accuracy
Broncho-	Group of complete lung (n=386)	0.795	0.802	0.800
pneumonia	Group of close-up lung (n=387)	1	0.250	0.647
	Equally blended group (n=386)	0.933	0.640	0.772
	Equally blended group (n=773)	0.967	0.723	0.820
Interstitial	Group of complete lung (n=386)	0.856	0.670	0.753
pneumonia	Group of close-up lung (n=387)	0.356	0.984	0.818
	Equally blended group (n=386)	0.687	0.854	0.802
	Equally blended group (n=773)	0.758	0.944	0.874
Pleuritis	Group of complete lung (n=386)	0	1	0.824
	Group of close-up lung (n=387)	0	0.986	0.812
	Equally blended group (n=386)	0.295	0.997	0.869
	Total blended group (n=773)	0.636	0.989	0.928
Pleuro-	Group of complete lung (n=386)	N/A	N/A	N/A
pneumonia	Group of close-up lung (n=387)	N/A	N/A	N/A
	Equally blended group (n=386)	N/A	N/A	N/A
	Total blended group (n=773)	0.190	1.000	0.958

Each value is average of the results of five tests with randomly selected testing subset; N/A: non-available.

Table 3. Evaluation of the classification depending on the number of images in model

	Number of training data	Sensitivity	Specificity	Accuracy
Broncho-	A quarter (n = 193)	0.924	0.355	0.625
pneumonia	A half (n =386)	0.933	0.640	0.772
	Three quarters (n = 580)	0.945	0.638	0.774
	Total (n = 773)	0.967	0.723	0.820
Interstitial	A quarter (n = 193)	0.452	0.925	0.797
pneumonia	A half (n =386)	0.687	0.854	0.802
	Three quarters (n = 580)	0.693	0.916	0.847
	Total (n = 773)	0.758	0.944	0.874
Pleuritis	A quarter (n = 193)	0.184	0.991	0.858
	A half (n =386)	0.295	0.997	0.869
	Three quarters (n = 580)	0.466	0.994	0.895
	Total (n = 773)	0.636	0.989	0.928
Pleuro-	A quarter (n = 193)	N/A	N/A	N/A
pneumonia	A half (n =386)	N/A	N/A	N/A
	Three quarters (n = 580)	0.160	0.998	0.953
	Total (n = 773)	0.190	1.000	0.958

Each value is average of the results of five tests with randomly selected testing subset; N/A: non-available.

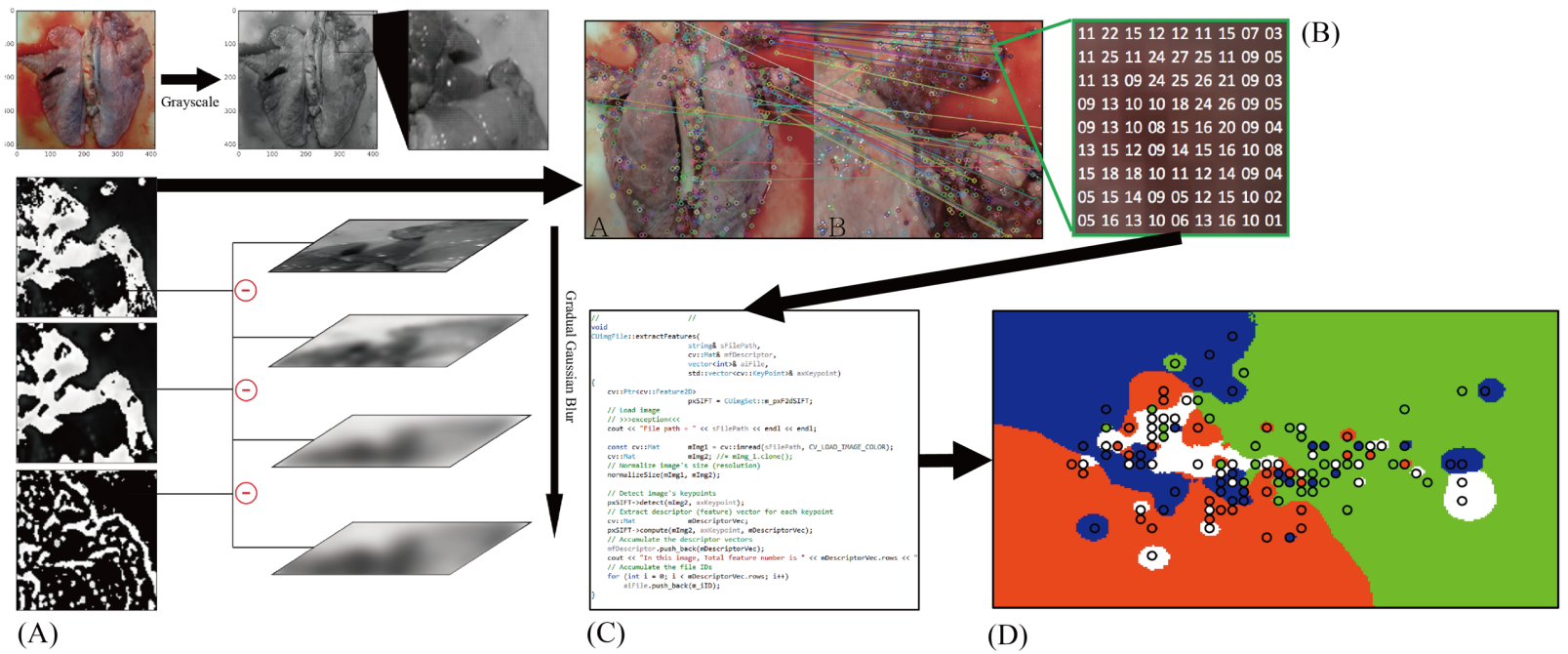

Fig. 1. Visualization of scale invariant feature transform (SIFT) extraction method and k-nearest neighbor classification. (A) scale-invariant feature transform (SIFT) extraction. (B) Example of extraction of SIFT features from gross image of swine lung lesion. (C) Computational recognition and analysis. (D) Conceptual visualization of KNN methodology classifying four classes of swine pneumonia

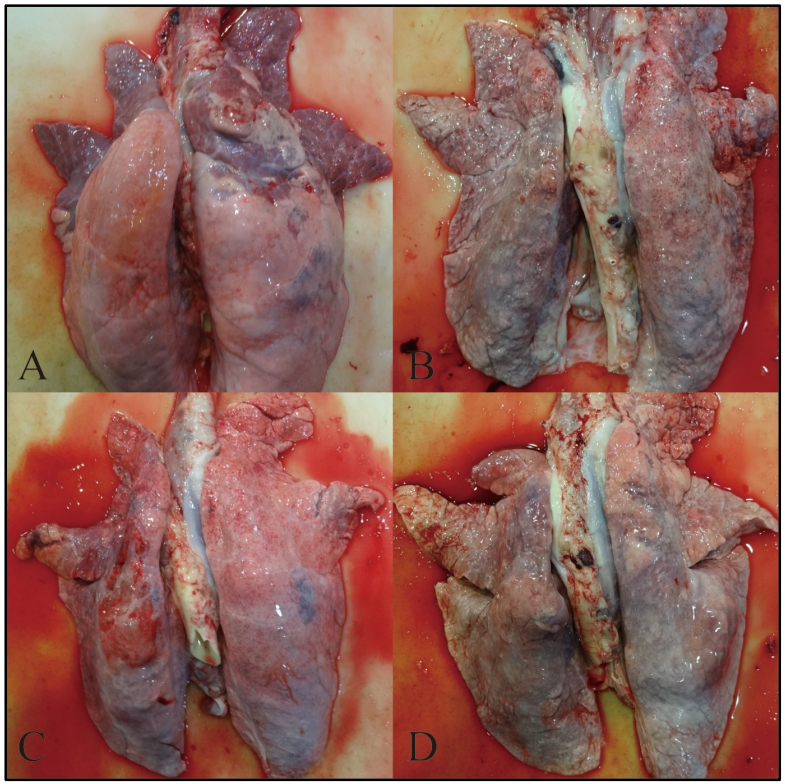

Fig. 2. Gross image of the lung diagnosed by histopathology and visual inspection; A: bronchopneumonia; B: interstitial pneumonia; C: pleuritis; D: pleuropneumonia

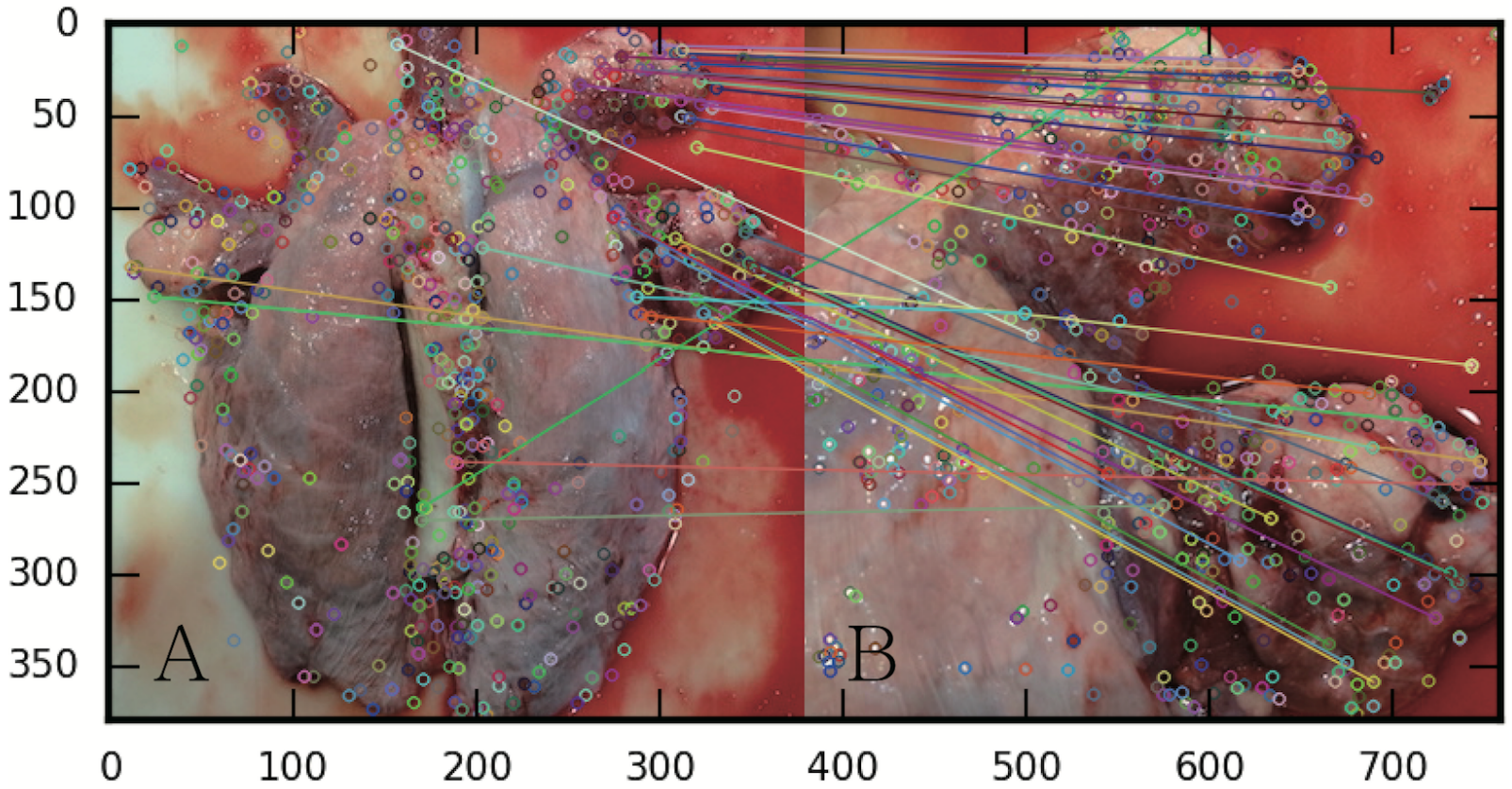

Fig. 3. Scale invariant feature transform features were extracted from gross image of swine lung. Each extracted feature is colored circle on image, and the couples of neighbors feature were connected with line.

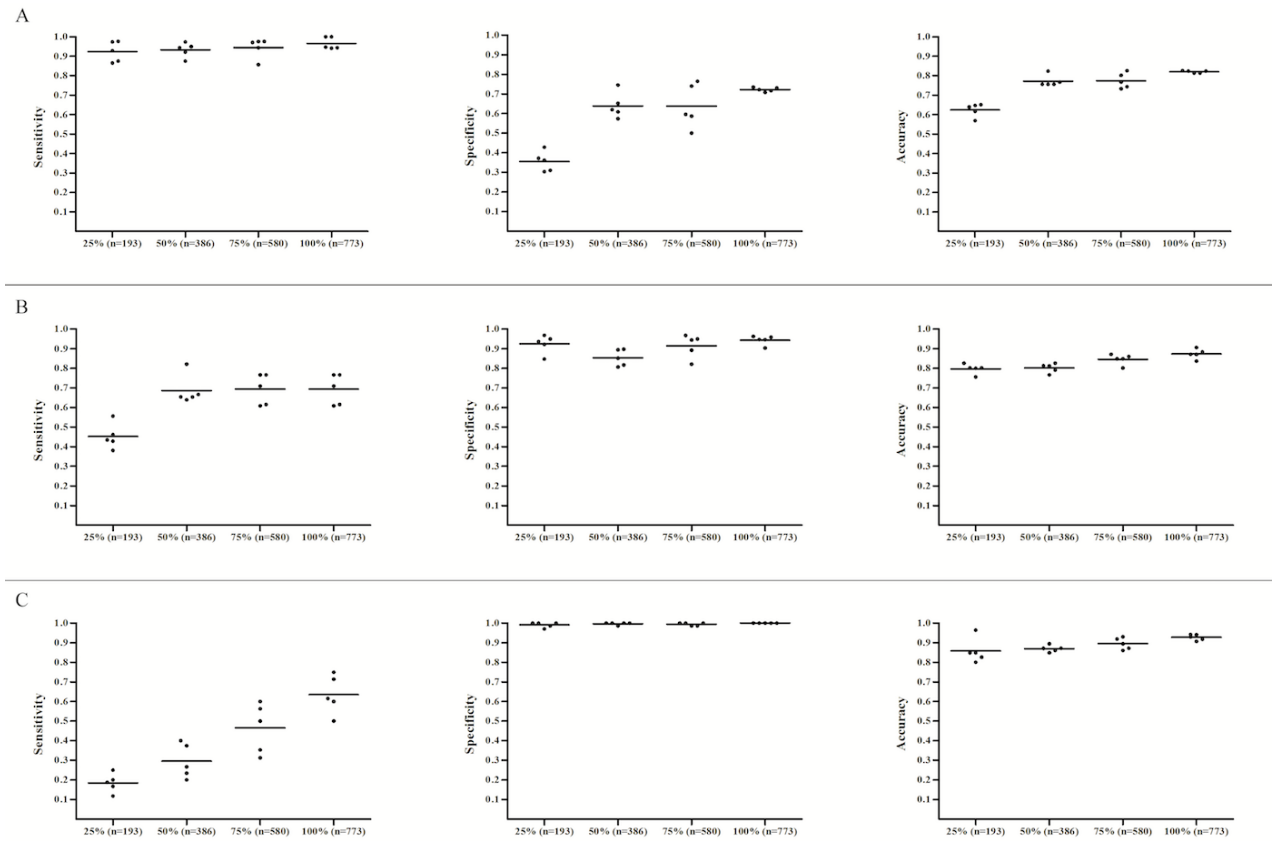

Fig. 4. Sensitivity, specificity and accuracy changes in the classification of (A) bronchopneumonia, (B) interstitial pneumonia, and (C) pleuritis according to the number of images in the model. Sensitivity is high for bronchopneumonia, specificity is high for interstitial pneumonia and pleuritis. Overall, increases in the indicators of diagnostic performance were observed depending on the number of trained data.

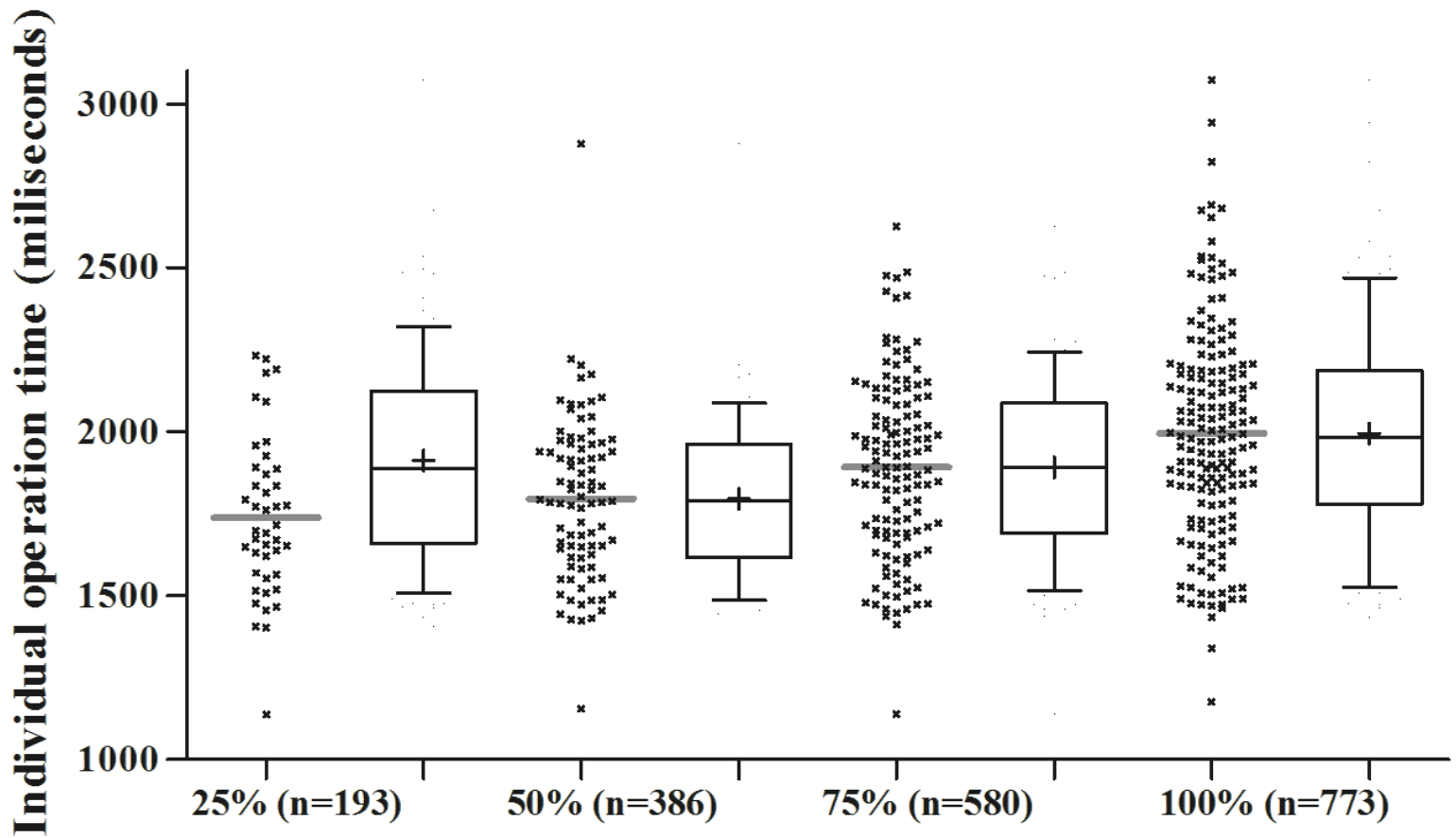

Fig. 5. Individual operation time for classification of one test image depending on the number of images in the model. As the total number of data increases, the runtime tends to increase.

Discussion

In the present study, CAD model for diagnosis of swine lung was established, and showed good diagnostic performance for bronchopneumonia and interstitial pneumonia. In order to estimate classification performance for training image, the diagnostic indicators were examined according to the characteristics of training image. As a result, a training set with mixture of complete and close-up lung image showed better performance to detect pleuritis in same number of image, but still showed low performance for pleuropneumonia.

CAD has been developed on field of diagnostic imaging such as radiography, computed tomography, and magnetic resonance imaging, using image detection techniques of computer vision (Freer and Ulissey, 2001). However, it was scarce that CAD study about pathologic diagnosis was investigated with image analysis of gross image which is obtained by common digital camera. In the past studies, visual inspection though the two-dimensional information of the photographed gross images showed high correlation with visual inspection for the three-dimensional real swine lung (Davies et al., 1995; Hurnik et al., 1993). Therefore, it was considered that the diagnostic model based on digital gross images is applicable to lung lesion monitoring in slaughter check system.

To find the better type of image to be more intensively collected in further studies, the classification performances of the CAD models were investigated according to image characteristics. In a visual inspection by naked eyes, it is possible to diagnose lung lesion with close-up images of partial lung. But if there is a gross image which

has all area of the lung, more accurate diagnosis is possible. Similar to the visual inspection, the CAD model showed better classification performance in the model based on complete lung image for diagnosis of bronchopneumonia and interstitial pneumonia. However, the model based on close-up lung images showed poor performance that was inapplicable to slaughter check system. It seemed that close-up images of partially area of lung would be not significant for establishment of CAD model. Nevertheless, in contrast that detection of pleuritis was unable in the both former two models, the model based on blended group which was integration of complete and close-up lung image showed availability to detect pleuritis. The influence of training image number was excluded by controlling the number of image in each group. Because the bag of feature model which is used in KNN classification integrates close-up and non-close-up feature (Wang et al., 2013), it is considered that collecting both of close-up and complete lung images is more effective to detect pleuritis than the collecting single type of image. Therefore, it is more desirable to collect integrated types of close-up and complete lung image than to collect only single type of images.

Except in pleuropneumonia, most diagnostic indicators showed positive correlation with the number of training images in Pearson correlation analysis (Cohen, 1988). Although sensitivity of bronchopneumonia, specificity of interstitial pneumonia, and specificity of pleuritis did not show significant correlation with the images quantity, the indicators were high enough. Though specificity of bronchopneumonia, sensitivity of interstitial pneumonia and

sensitivity of pleuritis is still inapplicable to slaughter check system (Table 3), it is expected that higher diagnostic performance would be achieved by collecting more image data because these diagnostic indicators are highly correlated with data quantity. To investigate the correlation between diagnostic indicator and data quantity, number of images in each disease is divided by quartiles. In the present study, however, the number of pleuropneumonia image was too small to construct enough size of testing subset and the performance changes of pleuropneumonia classification by increasing number of images could not be analyzed. Nonetheless, considering a positive correlation between sensitivity and quantity in pleuropneumonia, the classification performance of pleuropneumonia is also expected to improve if the data quantity of pleuropneumonia image increases.

The CAD model showed good performance in diagnosis of bronchopneumonia, interstitial pneumonia. The classification of bronchopneumonia in most models usually showed good sensitivity. Since bronchopneumonia detection was highly correlated between visual inspection and histopathology in past study (Ostanello et al., 2007), it is considered that classification of bronchopneumonia through CAD shows good performance to detect bronchopneumonia. Because the representative lesions of bronchopneumonia are usually located in cranial side of lung and have consolidated texture, it is considered that bronchopneumonia could be easily distinguished from the other pulmonary diseases using gross images. Considering that KNN classification is instance-based learning algorithm and the bronchopneumonia has the largest proportion of training images, the CAD model

could encounter overfitting to bronchopneumonia. However, because the CAD model showed fine diagnostic performance to interstitial pneumonia, it seems not to confront overfitting to bronchopneumonia.

In contrast to bronchopneumonia and interstitial pneumonia, classification of pleuritis were showed relatively low performance, and classification of pleuropneumonia was poor in the CAD model. The pleuritis and pleuropneumonia are pleural diseases which have visceral and parietal pleural adhesions, and fibrous surface as representative lesion. Since it is difficult to find out evidence of adhesion between visceral and parietal surface from photograph of separated lung, it is considered that more features which have label of pleural diseases were required to improve the performance. The number of pleuropneumonia samples was 49, even not enough to construct reliable testing subset. To secure minimal testing subset, more than 160 pleuropneumonia images are needed to attain at least four images as testing subset even if the group is divided by quartiles. In general, classification performance in machine learning is affected by the quantity and quality of the training data which construct background database for classification (Wei et al., 2013).

Although the CAD model of present study still warrant further study, and is not enough to be applied into industrial field, the positive correlation between classification performance and the number of training images was identified. Because KNN is instance based classification methods, it shows robustness which means the ability to withstand error data or noised training data. Because KNN is

also effective algorithm with numerous training data, the increasing pattern of the CAD model performance is considered to be continued as the training data increases (Kotsiantis et al., 2007). Therefore, if the increase pattern is maintained, excellent sensitivity and specificity would be expected. It is supposed that more than about 2000 samples of swine lung that is four-folded number of total images in this study would be required to establish CAD model to be applied to the practical stage in slaughter check system.

Low diagnostic performance of pleural disease is a concern of this study, but it is suggested that collecting both image types of complete and close-up lung is better to detect pleuritis than collecting only single image type of complete or close-up lung. KNN classification has high robustness, expecting excellent performance in larger size of images. And the positive correlation between sensitivity and data quantity is identified in pleuritis. Therefore, it is considered that additional collecting both image types could supplement the concern of pleural disease. The classification performance should be investigated in further study before applying in slaughter check system.

However, since KNN calculates every distances from numerous existing trained features to a new feature of the testing image, the time required for the individual classification increases as the number of image data in the model increases. Even though the performance of the classification model improves according to the number of training images, the individual operation time increases (Zhang et al., 2017). Therefore, when constructing a model from more massive training database,

high-level of computer system specification would be required to operate individual classification in real time at the slaughterhouses. If high-level of computer system specification is insufficient to reduce the individual operation time, it could be necessary to compromise operation time and diagnostic performance for real-time application to the slaughter check system in further study.

The CAD model showed good classification performance for bronchopneumonia and interstitial pneumonia, suggesting applicability of CAD model to slaughterhouses as pulmonary lesion monitoring system. Although low diagnostic performance of pleural disease is a concern of this study, diagnostic performances are considered to be improved as the number of training data increases. In addition, rather than collecting only single type of images which is complete or close-up lung, integrating two types of the model would helpful to detect pleuritis. This is a novel approach that applies machine learning based CAD to the veterinary diagnostic field, especially to gross pathology. It is expected to be an efficient decision supporting tool for organ inspection of pigs as well as other livestock animals in veterinary field, and also contribute to disease management of farms, and food hygiene.

Conclusion

In the present study, analysis of correlation between visual lung score and histopathological diagnosis was revealed the visual lung score could be a screening test for bronchopneumonia with excellent discrimination. It is considered that the visual information of the photograph is valid to diagnosis, so that the gross lung images was subjected to machine learning algorithm to establish CAD model for swine lung. The CAD model showed applicability to slaughterhouses for pulmonary lesion monitoring system. It still warrants further study to improve classification performance, however, the analysis shows that better classification performance of CAD models for bronchopneumonia, interstitial pneumonia, and pleuritis is expected if more training images are available.

References

Altman, N. S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression." Am Stat 46(3): 175-185.

Aly, A. A., S. B. Deris and N. Zaki (2011). "Research review for digital image segmentation techniques." Int J Inf Tech 3(5): 99.

André, B., T. Vercauteren, A. Perchant, A. M. Buchner, M. B. Wallace and N. Ayache (2009). Endomicroscopic image retrieval and classification using invariant visual features. Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on, IEEE.

Awai, K., K. Murao, A. Ozawa, M. Komi, H. Hayakawa, S. Hori and Y. Nishimura (2004). "Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance." Radiology 230(2): 347-352.

Baptista, F., J. Dahl and L. R. Nielsen (2010). "Factors influencing Salmonella carcass prevalence in Danish pig abattoirs." Prev Vet Med 95(3): 231-238.

Burroni, M., R. Corona, G. Dell'Eva, F. Sera, R. Bono, P. Puddu, R. Perotti, F. Nobile, L. Andreassi and P. Rubegni (2004). "Melanoma computer-aided diagnosis." Clin Cancer Res 10(6): 1881-1886.

Chae, C. (2005). "A review of porcine circovirus 2-associated syndromes and diseases." Vet J 169(3): 326-336.

Choi, Y. K., S. M. Goyal and H. S. Joo (2003). "Retrospective analysis of etiologic

agents associated with respiratory diseases in pigs." Can. Vet. J. 44(9): 735-737.

Christensen, J., B. Ellegaard, B. K. Petersen, P. Willeberg and J. Mousing (1994). "Pig health and production surveillance in Denmark: sampling design, data recording, and measures of disease frequency." Prev. Vet. Med. 20(1-2): 47-61.

Cohen, J. (1988). "Statistical power analysis for the behavioral sciences . Hilsdale." NJ: LEA 2: 79-83.

Danielsson, P.-E. (1980). "Euclidean distance mapping." Comput Gr Image Process 14(3): 227-248.

Davies, P., P. Bahnson, J. Grass, W. Marsh and G. Dial (1995). "Comparison of methods for measurement of enzootic pneumonia lesions in pigs." Am J Vet Res 56(6): 9-14.

Doi, K. (2007). "Computer-aided diagnosis in medical imaging: historical review, current status and future potential." Comput Med Imaging Graph 31(4): 198-211.

El-Dahshan, E.-S. A., H. M. Mohsen, K. Revett and A.-B. M. Salem (2014). "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm." Expert Syst Appl 41(11): 5526-5545.

Elbers, A., J. Braamskamp, L. Dekkers, R. Voets, T. Duinhof, W. Hunneman and J. Stegeman (2000). "Aujeszky's disease virus eradication campaign successfully heading for last stage in the Netherlands." Vet Q 22(2): 103-

107.

Feng, Z.-X., G.-Q. Shao, M.-J. Liu, X.-S. Wu, Y.-Q. Zhou and G. Yuan (2010). "Immune responses to the attenuated Mycoplasma hyopneumoniae 168 strain vaccine by intrapulmonic immunization in piglets." Agric. Sci. China 9(3): 423-431.

Fernández-Delgado, M., E. Cernadas, S. Barro and D. Amorim (2014). "Do we need hundreds of classifiers to solve real world classification problems." J Mach Learn Res 15(1): 3133-3181.

Freer, T. W. and M. J. Ulissey (2001). "Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center." Radiology 220(3): 781-786.

Gardner, I. A., P. Willeberg and J. Mousing (2002). "Empirical and theoretical evidence for herd size as a risk factor for swine diseases." Anim. Health Res. Rev. 3(01): 43-55.

Gurcan, M. N., B. Sahiner, N. Petrick, H. P. Chan, E. A. Kazerooni, P. N. Cascade and L. Hadjiiski (2002). "Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system." Med Phys 29(11): 2552-2558.

Hansen, M. S., S. E. Pors, H. Jensen, V. Bille-Hansen, M. Bisgaard, E. M. Flachs and O. L. Nielsen (2010). "An investigation of the pathology and pathogens associated with porcine respiratory disease complex in Denmark." J. Comp. Pathol. 143(2): 120-131.

Harms, P. A., P. G. Halbur and S. D. Sorden (2002). "Three cases of porcine respiratory disease complex associated with porcine circovirus type 2 infection." J. Swine Health Prod. 10(1): 27-30.

Hawkins, D. M. (2004). "The problem of overfitting." J Chem Inf Comput Sci 44(1): 1-12.

Hosmer Jr, D. W., S. Lemeshow and R. X. Sturdivant (2013). Applied logistic regression, John Wiley & Sons.

Huhn, R. (1970). "Swine enzootic pneumonia: incidence and effect on rate of body weight gain." American journal of veterinary research 31: 1097-1108.

Hurnik, D., I. R. Dohoo and L. A. Bate (1994). "Types of farm management as risk factors for swine respiratory disease." Prev. Vet. Med. 20(1-2): 147-157.

Hurnik, D., P. E. Hanna and I. R. Dohoo (1993). "Evaluation of rapid gross visual appraisal of swine lungs at slaughter as a diagnostic screen for enzootic pneumonia." Can J Vet Res 57(1): 37-41.

Hwang, W. M. and J. H. Han (2006). "Prevalence of endemic diseases and effect on control of respiratory diseases in pig farms through slaughter check and management changes." Korean J Vet. Public Health 30(1): 27-56.

Jäger, H. C., T. J. McKinley, J. L. Wood, G. P. Pearce, S. Williamson, B. Strugnell, S. Done, H. Habernoll, A. Palzer and A. W. Tucker (2012). "Factors associated with pleurisy in pigs: a case-control analysis of slaughter pig data for England and Wales." PLoS One 7(2): e29655.

Jeffrey, J., A. Locke and J. Kent (2012). Disease of swine, USA, John Wiley&Sons,

Inc.

Jiang, M., S. Zhang, H. Li and D. N. Metaxas (2015). "Computer-aided diagnosis of mammographic masses using scalable image retrieval." IEEE Trans Biomed Eng 62(2): 783-792.

Joglekar, J., S. S. Gedam and B. K. Mohan (2014). "Image matching using SIFT features and relaxation labeling technique—A constraint initializing method for dense stereo matching." IEEE Trans Geosci Remote Sens 52(9): 5643-5652.

Kakeda, S., J. Moriya, H. Sato, T. Aoki, H. Watanabe, H. Nakata, N. Oda, S. Katsuragawa, K. Yamamoto and K. Doi (2004). "Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system." AJR Am J Roentgenol 182(2): 505-510.

Kashif, M., T. M. Deserno, D. Haak and S. Jonas (2016). "Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment." Comput Biol Med 68: 67-75.

Kelly, J. G., P. P. Angelov, J. Trevisan, A. Vlachopoulou, E. Paraskevaidis, P. L. Martin-Hirsch and F. L. Martin (2010). "Robust classification of low-grade cervical cytology following analysis with ATR-FTIR spectroscopy and subsequent application of self-learning classifier eClass." Anal Bioanal Chem 398(5): 2191-2201.

Khanmohammadi, M., R. Nasiri, K. Ghasemi, S. Samani and A. B. Garmarudi (2007). "Diagnosis of basal cell carcinoma by infrared spectroscopy of

whole blood samples applying soft independent modeling class analogy." J Cancer Res Clin Oncol. 133(12): 1001-1010.

Kim, N.-H., W.-M. Hwang, J.-G. Lee, S.-M. Lee, D.-S. Yang, C.-H. Lee, S.-J. Kim and J.-H. Han (2011). "Pathological observation of porcine respiratory disease in slaughter pigs." Korean J Vet Serv. 34(4): 389-395.

Kotsiantis, S. B., I. Zaharakis and P. Pintelas (2007). Supervised machine learning: A review of classification techniques.

Lee, C.-H., W.-M. Hwang, J.-G. Lee, S.-M. Lee, S.-J. Kim, N.-H. Kim, D.-S. Yang and J.-H. Han (2011). "Study on gross finding of lung lesions and causative pathogens of porcine respiratory disease complex from slaughtered pigs in Incheon." Korean J Vet Serv. 34(4): 313-320.

Lee, H. and Y.-P. P. Chen (2015). "Image based computer aided diagnosis system for cancer detection." Expert Syst Appl 42(12): 5356-5365.

Liedlgruber, M. and A. Uhl (2011). "Computer-aided decision support systems for endoscopy in the gastrointestinal tract: a review." IEEE Rev Biomed Eng 4: 73-88.

Lison, P. (2015). “An introduction to machine learning.

Lowe, D. G. (2004). "Distinctive image features from scale-invariant keypoints." Int J Comput Vis. 60(2): 91-110.

Madec, F. and M. Kobisch (1982). "Bilan lésionnel des poumons de porcs charcutiers à l’abattoir." J Rech. Porcine. 14: 405-412.

Maes, D., H. Deluyker, M. Verdonck, F. Castryck, C. Miry, B. Vrijens and A. d.

Kruif (1999). "Risk indicators for the seroprevalence of Mycoplasma hyopneumoniae, porcine influenza viruses and Aujeszky's disease virus in slaughter pigs from fattening pig herds." Zoonoses Public Health 46(5): 341-352.

Meyns, T., J. Van Steelant, E. Rolly, J. Dewulf, F. Haesebrouck and D. Maes (2011). "A cross-sectional study of risk factors associated with pulmonary lesions in pigs at slaughter." Vet J 187(3): 388-392.

Morrison, R. B., H. D. Hilley and A. D. Leman (1985). "Comparison of methods for assessing the prevalence and extent of pneumonia in market weight swine." Can Vet J 26(12): 381.

Mousing, J., P. T. Jensen, C. Halgaard, F. Bager, N. Feld, B. Nielsen, J. Nielsen and S. Bech-Nielsen (1997). "Nation-wide Salmonella enterica surveillance and control in Danish slaughter swine herds." Prev. Vet. Med. 29(4): 247-261.

Muja, M. and D. Lowe (2009) "Flann-fast library for approximate nearest neighbors user manual."

Neumann, E. J., J. B. Kliebenstein, C. D. Johnson, J. W. Mabry, E. J. Bush, A. H. Seitzinger, A. L. Green and J. J. Zimmerman (2005). "Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States." J Am Vet Med Assoc 227(3): 385-392.

Ostanello, F., M. Dottori, C. Gusmara, G. Leotti and V. Sala (2007). "Pneumonia

Disease Assessment using a Slaughterhouse Lung-Scoring Method." Transbound Emerg Dis 54(2): 70-75.

Ostanello, F., M. Dottori, C. Gusmara, G. Leotti and V. Sala (2007). "Pneumonia Disease Assessment using a Slaughterhouse Lung-Scoring Method." J Vet Med A 54(2): 70-75.

Poteet, B. A. (2008). "Veterinary teleradiology." Vet Radiol Ultrasound 49(s1).

Redondo, E., A. Masot, A. Fernandez and A. Gazquez (2009). "Histopathological and immunohistochemical findings in the lungs of pigs infected experimentally with Mycoplasma hyopneumoniae." J. Comp. Pathol. 140(4): 260-270.

Robertson, J., D. Wilson and W. Smith (1990). "Atrophic rhinitis: The influence of the aerial environment." Anim. Sci. 50(1): 173-182.

Rueda, D., C. Bulnes, R. Durand and P. Bustamante (2002). "Morphological evaluation of porcine pneumonias in slaughterhouses by using a score method." Rev. Salud. Anim. 24(3): 208-212.

Samuel, A. L. (2000). "Some studies in machine learning using the game of checkers." IBM J Res Dev 44(1.2): 206-226.

Sanchez-Vazquez, M., W. Strachan, D. Armstrong, M. Nielen and G. Gunn (2011). "The British pig health schemes: integrated systems for large-scale pig abattoir lesion monitoring." Vet Rec: 10.1136/vr. d4814.

Schapire, R. E. (2003). The boosting approach to machine learning: An overview.

Nonlinear estimation and classification, Springer: 149-171.

Sethian, J. A. (1999). Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science, Cambridge university press.

Setio, A. A. A., F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S. J. van Riel, M. M. W. Wille, M. Naqibullah, C. I. Sánchez and B. van Ginneken (2016). "Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks." IEEE Trans Med Imag 35(5): 1160-1169.

Shouman, M., T. Turner and R. Stocker (2012). "Applying k-nearest neighbour in diagnosing heart disease patients." Int J Inf Educ Technol 2(3): 220.

Sibila, M., M. Pieters, T. Molitor, D. Maes, F. Haesebrouck and J. Segalés (2009). "Current perspectives on the diagnosis and epidemiology of Mycoplasma hyopneumoniae infection." Vet. J. 181(3): 221-231.

Sokolova, M. and G. Lapalme (2009). "A systematic analysis of performance measures for classification tasks." Inf Process Manag 45(4): 427-437.

Stärk, K., D. Pfeiffer and R. Morris (1998). "Risk factors for respiratory diseases in New Zealand pig herds." N. Z. Vet. J. 46(1): 3-10.

Stärk, K. D. (2000). "Epidemiological investigation of the influence of environmental risk factors on respiratory diseases in swine—a literature review." Vet. J. 159(1): 37-56.

Strawn, G. and C. Strawn (2015). "Moore's law at fifty." IT Professional 17(6): 69-

72.

Sun, T., J. Wang, X. Li, P. Lv, F. Liu, Y. Luo, Q. Gao, H. Zhu and X. Guo (2013). "Comparative evaluation of support vector machines for computer aided diagnosis of lung cancer in CT based on a multi-dimensional data set." Comput Methods Programs Biomed 111(2): 519-524.

Thacker, E. L. (2004). "Diagnosis of Mycoplasma hyopneumoniae." Anim. Health Res. Rev. 5(02): 317-320.

Thacker, E. L., P. G. Halbur, R. F. Ross, R. Thanawongnuwech and B. J. Thacker (1999). "Mycoplasma hyopneumoniae potentiation of porcine reproductive and respiratory syndrome virus-induced pneumonia." J. Clin. Microbiol. 37(3): 620-627.

Tsomko, E., H. Kim and E. Izquierdo (2010). "Linear Gaussian blur evolution for detection of blurry images." IET Image Processing 4(4): 302-312.

van Ginneken, B., C. M. Schaefer-Prokop and M. Prokop (2011). "Computer-aided diagnosis: how to move from the laboratory to the clinic." Radiology 261(3): 719-732.

Veta, M., J. P. Pluim, P. J. Van Diest and M. A. Viergever (2014). "Breast cancer histopathology image analysis: A review." IEEE Trans Biomed Eng 61(5): 1400-1411.

Vicca, J., T. Stakenborg, D. Maes, P. Butaye, J. Peeters, A. de Kruif and F. Haesebrouck (2003). "Evaluation of virulence of Mycoplasma hyopneumoniae field isolates." Vet. Microbiol. 97(3): 177-190.

Wang, J. J.-Y., H. Bensmail and X. Gao (2013). "Joint learning and weighting of visual vocabulary for bag-of-feature based tissue classification." Pattern Recognit 46(12): 3249-3255.

Wei, Z., W. Wang, J. Bradfield, J. Li, C. Cardinale, E. Frackelton, C. Kim, F. Mentch, K. Van Steen and P. M. Visscher (2013). "Large sample size, wide variant spectrum, and advanced machine-learning technique boost risk prediction for inflammatory bowel disease." Am J Hum Genet 92(6): 1008-1012.

Wilson, M., R. Takov, R. Friendship, S. Martin, I. McMillan, R. Hacker and S. Swaminathan (1986). "Prevalence of respiratory diseases and their association with growth rate and space in randomly selected swine herds." Can. J. Vet. Res. 50(2): 209-216.

Zhang, S., X. Li, M. Zong, X. Zhu and R. Wang (2017). "Efficient kNN Classification With Different Numbers of Nearest Neighbors." IEEE Trans. Neural Netw. Learn. Syst.

국문초록

영상 기반 기계학습을 이용한 돼지 폐병변의 진단

이 홍 석

지도 교수: 김 용 백

서울대학교 대학원 수의학과 임상수의학 (임상병리학) 전공

양돈 선진국에서 도체검사는 식품 위생 및 돼지 질병 모니터링 및 방역 계획 수립에 사용되고 있다. 현재 임상 진단 영역에서 적용되고 있는 기계학습 기법 기반의 컴퓨터 진단을 도체검사에 적용시킨다면 폐, 간, 장 등의 장기 검사를 보다 개선 시킬 수 있을 것으로 기대된다. 본 연구에서는 영상 기반 진단 시스템 구축을 위한 선행연구로써 돼지 폐의 사진을 통해 평가한 폐병변지수와 기관지성 폐렴과의 상관성이 조사되었다. 폐 조직과 폐병변 이미지는 도축장에서 수집되었다. 폐병변를 통해 폐뼝변지수가 평가되었으며 조직병리학적 검사를 통해 폐병변은

기관지성폐렴과 간질성 폐렴으로 구분되었다. 폐병변 지수는 90% 신뢰구간에서 기관지성 폐렴에 대해 100%의 민감도와 77.3%의 특이도를 지니는 것으로 확인되었다. 그러나 기관지성폐렴의 진단에 대해서는 특이도가 낮았다. 수신자 조작특성 곡선 아래 면적은 0.896 으로, 폐병변 지수는 기관지성 폐렴에 대한 감별력이 우수한 것으로 판단되었다. 질병 폐의 사진을 통한 시각정보가 스크리닝 테스트에 유용하게 작용하였으므로, 폐 사진을 학습 데이터집합으로 구성하여 영상 기반 기계학습을 통한 돼지 폐병변 진단 모델을 수립하였다. 폐병변 이미지와 폐 조직은 도축장으로부터 수집되었으며, 폐 병변은 조직병리학적으로 기관지성 폐렴, 간질성 폐렴, 그리고 흉막성 질환들로 분류되었다. 진단 모델을 구성함에 있어 사진의 주요 특징점을 추출하는 기법으로 *scale-invariant feature transform* 이 적용되었으며, 추출된 특징점을 분류하기 위한 기계학습 분류 기법으로 *k-nearest neighbor* 이 적용되었다. 이미지 기반의 진단 모델은 기관지성 폐렴에 대해서 96.7%의 높은 민감도와 72.3%의 특이도, 그리고 82.0%의 정확도를 보였으며, 간질성 폐렴에 대해서는 9.4%의 높은 특이도와 87.4%의 정확도, 그리고 75.8%의 민감도를 보였다. 그러나 흉막염 및 흉막 폐렴의 진단에 있어서는 비교적 낮은 성능이 확인되었다. 본 연구는 도축장에서의 도체검사에 컴퓨터 보조진단의 적용 가능성을 타진하는, 영상 기반 기계학습을 이

용한 장기검사에 대한 새로운 접근법을 제시하였다. 본 연구에서 얻어진 정보들은 수의 진단 영역에서 영상을 기반으로 한 컴퓨터 도체검사의 초석이 될 것으로 기대된다.

핵심어: 컴퓨터 보조 진단, 기계학습, 조직병리학, 폐렴, 도체검사, 폐병변 지수

학번 : 2017-21782